

Seq primer: -21M13 Forward
High quality sequence stop: 673
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1. 673
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:H3102B03-3"
/db_xref="taxon:10090"
/clone="H3102B03"
/clone_1b="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 137 a 191 c 173 g 172 t
ORIGIN

Query Match 50.6%; Score 646.2; DB 12; Length 673;
Best Local Similarity 99.3%; Pred. No. 7.1e-103;
Matches 670; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 585 CGAGGAGGGCGATGTTTCAGGCTGCAGGCGCGCGGGTGGAGCAGCACCACCGGA 644
Db 673 CGAGGAGGGCGATGTTTCAGGCTGCAGGCGCGCGGGTGGAGCAGCACCACCGGA 614

QY 645 GGACCGGAGACAGTGGCGCGGATGCAGTCTGAGCCTGGAGCGAGGCCATGTCCTGC 704
Db 613 GGACCGGAGACAGTGGCGCGGATGCAGTCTGAGCCTGGAGCGAGGCCATGTCCTGC 554

QY 705 CGCAGAGATGGCTCAGAGACCCCGGTGATGCGATGCCCTCGAGACCGAGCCCTCCCGCA 764
Db 553 CGCAGAGATGGCTCAGAGACCCCGGTGATGCGATGCCCTCGAGACCGAGCCCTCCCGCA 494

QY 765 AAGCAGGAGCAGAGAGAGCGCGCTGCTCCAGTCTTAGAGCAGAGTAGCGGCTA 824
Db 493 AAGCAGGAGCAGAGAGAGCGCGCTGCTCCAGTCTTAGAGCAGAGTAGCGGCTA 434

QY 825 CTATCACTGCAAGACTGCAAAATCCGGTGGAGAGCGCCTATGTTGTTGTGTCAGGG 884
Db 433 CTATCACTGCAAGACTGCAAAATCCGGTGGAGAGCGCCTATGTTGTTGTGTCAGGG 374

QY 885 CACCAGTAAGGTGTTACTTCAACAGTCTGCCAGTGTGTGAGAAATCCTACCAACCTT 944
Db 373 CACCAGTAAGGTG-TACTTCAACAGTCTGCCAGTGTGTGAGAAATCCTACCAACCTT 315

QY 945 ACAGAGTGGAGGACATCAGTCTCAAGTGTGTAAGAAGTATGATGTGCTGCCAGTCA 1004
Db 314 ACAGAGTGGAGGACATCAGTCTCAAGTGTGTAAGAAGTATGATGTGCTGCCAGTCA 255

QY 1005 GATTTCGCCACGTGAGACCTTAACAGCCCATCGGCAAGACTTGTGTGGAGATGCAAG 1064
Db 254 GATTTCGCCACGTGAGACCTTAACAGCCCATCGGCAAGACTTGTGTGGAGATGCAAG 195

QY 1065 ACAAGGCTGCTCTGCGACAGCAGCCTTCAGCTTCAATATCATTTAGTACAGTCA 1124
Db 194 ACAAGGCTGCTCTGCGACAGCAGCCTTCAGCTTCAATATCATTTAGTACAGTCA 135

QY 1125 AAACGTTTCTGCTAGATGGGCTTAATGGAATGACACAGTACCTTCTCCCTCTTCACC 1184
Db 134 AAACGTTTCTGCTAGATGGGCTTAATGGAATGACACAGTACCTTCTCCCTCTTCACC 75

QY 1185 TCTTCCCTTCCAAATCTTTCATGACAGACAGTGTACTTGTGATTAAGCCTGTGAATA 1244
Db 74 TCTTCCCTTCCAAATCTTTCATGACAGACAGTGTACTTGTGATTAAGCCTGTGAATA 16

QY 1245 AAAGTATTCGAAC 1259
Db 15 AAAGTATTCGAAC 1

RESULT 2
BB641267 675 bp mRNA linear EST 26-OCT-2001
DEFINITION BB641267 RIKEN full-length enriched, 10 days neonate cortex Mus
LOCUS BB641267
ACCESSION BB641267
VERSION BB641267.1 GI:16476392
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE
1 (bases 1 to 675)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
TITLE Yoshinide Hayashizaki
JOURNAL Contact: Yoshinide Hayashizaki
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,M., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagl,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,I., Alizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

FEATURES
source
Location/Qualifiers
1. 675
/organism="Mus musculus"

BASE COUNT	130	a	204	c	236	g	100	t	5 others
ORIGIN									

/note_site_1: Salt; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGCAGAGAGGATCCACAGAGCTCTTTT TTTTTTTTTT VN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGCAGAGAGATTCTCGAGGTAAATTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda

FLC I."

Query Match	50.4%;	Score 643;	DB 10;	Length 675;
Best Local Similarity	97.68;	Pred. No. 2.5e-102;		
Matches 660; Conservative	0;	Mismatches 15;	Indels 1;	Gaps 1

QY	286	TCGGTGCAGTGTTCAC	TCTCGGGCGCCGACAGCCTGCAGGGTCCGAGCCAGCCCC	345
Db	1	TGGGTGCAGTGTTCAC	TCTCGGGCGCCGACAGCCTGCAGGGTCCGAGCCAGCCCC	60
QY	346	GACGCCCCGATCGGGTTC	CTGTCAACCCCGCTGGCCACGCGCGCGCGGAGATCCCCGGA	405
Db	61	GACGCCCCGATCGGGTTC	CTGTCAACCCCGCTGGCCACGCGCGCGCGGAGATCCCCGGA	120
QY	406	TCCTGGCAGACCCGTAG	CCCCGTTCTCTGCTCCGTGACCTTCTGTGGCCCTCTCCTCACTG	465
Db	121	TCCTGGCAGACCCGTAG	CCCCGTTCTCTGCTCCGTGACCTTCTGTGGCCCTCTCCTCACTG	180
QY	466	GAGGTTCGCGGAGCCAG	GCCAGACACCCACGAAGGAGAGGGAGCGCCGATCCTCGGGG	525
Db	181	GAGGTTCGCGGAGCCAG	GCCAGACACCCACGAAGGAGAGGGAGCGCCGATCCTCGGGG	240
QY	526	ACCCGGGAACCCGGAAG	CCGAGAGAGGTGGCCCGCGAGGAAAGCGGTCCCCAGCCGGAAGC	585
Db	241	ACCCGGGAACCCGGAAG	CCGAGAGAGGTGGCCCGCGAGGAAAGCGGTCCCCAGCCGGAAGC	300
QY	586	GAGGAGGGCGATGTTCA	GCGCTGCAGGGGCAAGCCGGGTGGGACAGCACACCACCGGAG	645
Db	301	GAGGAGGGCGATGTTCA	GCGCTGCAGGGGCAAGCCGGGTGGGACAGCACACCACCGGAG	360
QY	646	GACCGGAACAGTGTGG	CGCGCGCATGCACTGAGCCTGGGACGACGAGCCATGTCTGCC	705
Db	361	GACCGGAACAGTGTGG	CGCGCGCATGCACTGAGCCTGGGACGACGAGCCATGTCTGCC	420
QY	706	GCAGAGATGGCTCAGGA	CCCCCGGTGATTGCGATGCCCTCGAGACCAAGGCTCCCGCA	765
Db	421	GCAGAGATGGCTCAGGA	CCCCCGGTGATTGCGATGCCCTCGAGACCAAGGCTCCCGCA	480
QY	766	AGCACGGAGCAGGACA	GAGCGCGCTGCTTTCAGTTCTTAGAGCAGAAGTACGGCTAC	825
Db	481	AGCACGGAGCAGGACA	GAGCGCGCTGCTTTCAGTTCTTAGAGCAGAAGTACGGCTAC	540
QY	826	TATCACTGCAAGGACT	GCAGAAATCCGGTGGGAGACGCGCTATGTGTGTGTGACAGGGC	885
Db	541	TATCANCTGCAGGACT	GCAGAAATCCGGTGGGAGACGCGCTATGTGTGTGTGACAGGGC	600
QY	886	ACCAAGTAGGTGTACT	TCAACACAGTTCTGCGCAGTGTGTGAGAAATCTTACACCCCTTA	945

Db 601 ACCAGTINAGTTC-TACTTCAAACANGTCTGCCGAGTGTGTGAGAAATCNCCTACACCCTTA 659

QY 946 CAGAGTGGAGCAGCATTC 961
|||||
|||||

Db 660 CAGAGTGGAGGAGCATTC 675

RESULT	3			
LOCUS	BB264222			
DEFINITION	BB264222	666 bp	mRNA	linear
ACCESSION	BB264222	RIKEN full-length enriched, 10 days neonate cortex Mus	EST	24-OCT-2001
VERSION	BB264222	musculus cDNA clone A830014H23 3', mRNA sequence.		
KEYWORDS	BB264222.2	GI:16400187		
SOURCE	EST.			
ORGANISM	house mouse.			
	Mus musculus			

REFERENCE
AUTHORS

TITLE _____
JOURNAL _____
COMMENT _____

On ~~JUL 17, 2000~~ this sequence version replaced gl:8960678.
Contact: Yoshihide Hayashizaki

RIKEN Mouse ESTs (Arakawa, T., et al. 2001).

Muramatsu, M. and Hayashizaki, Y.

Tagami, M., Tagawa, A., Takanashi, F., Takeda, Y., Tanaka, T.,
Muramatsu, M. and Hayashizaki, Y. Toya, T.,

Tagami, M., Tagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T.,
Tagami, M., Tagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,

OKAZAKI, Y., OKIDO, T., SATO, R., SAKAI, C., SAKAI, K., SANO, H., SASAKI, D., SHIBATA, K., SHINAGAWA, Y., CHIRAKI, T.

M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.

Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,T., Kobayashi,H.,
Kumada,S., Furuno,M., Hanagaki,T., Hara,A.

1 (bases 1 to 666)
Arakawa, T. Carninci P. Eukarya c. Drosophila m. H. sapiens

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

house mouse.
Mus musculus

EST.

BB264222.2 GI:16400187

musculus cDNA clone A830014H23 3', mRNA sequence.
BR264222

BB264222 RIKEN full-length enriched, 10 days neonate cortex Mus

BB264222 666 bp mRNA 1100bp 24 cm 20

1

1

[illegible]

AGAGTGGAGGACATTC 675

2AGAGTGGAGGACATC 961

...TCTGAGATGGTGACGAATCNCCTACACCCTTA 659

[illegible]

Page

1

БЕЛАТІДІ

Source

```
1. .666
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A830014H23"
```


QY 1248 GGTAATTGC 1255
 |||||
 Db 659 GGTAATTGC 666
 RESULT 4
 BG084538
 LOCUS
 DEFINITION
 BG084538 593 bp mRNA linear EST 26-JAN-2001
 H3102B03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3102B03 5', mRNA sequence.
 BG084538
 BG084538.1 GI:12567102
 EST.
 SOURCE
 house mouse.
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 593)
 Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
 T.S., Carter, M.G. and Ko, M.S.H.
 Verification and initial annotation of NIA mouse 15K cDNA clone set
 TITLE
 JOURNAL
 Unpublished (2001)
 COMMENT
 Other ESTs: H3102B03-3

1248 GGTATTGC 1255
11111111
659 GGTATTGC 666

RESULT 4
36084538
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM
SOURCE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

593 bp mRNA linear EST 26-JAN-2001
H3102B03-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
H3102B03 5', mRNA sequence.
BG084538
BG084538.1 GI:12567102
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 593)
Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
, T.S., Carter, M.G. and Ko, M.S.H.
Verification and initial annotation of NIA mouse 15K cDNA clone set
Unpublished (2001)
Other ESTs: H3102B03-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdnaelgusn.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3102 row: B column: 03
Seq primer: -21M13 Reverse
High quality sequence stop: 593
POLYA=No.

FEATURES
source

Location/Qualifiers
1..593
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:H3102B03-5"
/db_xref="taxon:10090"
/clone="H3102B03"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA
libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 15,247 clones from 11
embryo cDNA libraries (including preimplantation stage
embryos from unfertilized egg to blastocyst, embryonic
part of E7.5 embryos, extraembryonic part of E7.5 embryos
, and E12.5 female mesonephros/gonad) and one newborn
ovary cDNA library. Average insert size 1.5 kb. All
source libraries are cloned unidirectionally with Oligo(dT
)-Not primers. References include: (1) Genome-wide
expression profiling of mid-gestation placenta and embryo
using a 15,000 mouse developmental cDNA microarray, 2000,
Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
Large-scale cDNA analysis reveals phased gene expression
patterns during preimplantation mouse development, 2000,
Development, 127: 1737-1749; (3) Genome-wide mapping of
unselected transcripts from extraembryonic tissue of
7.5-day mouse embryos reveals enrichment in the t-complex
and under-representation on the x chromosome, 1998, Hum
Mol Genet 7: 1967-1978."

BASE COUNT 142 a 165 c 192 g 94 t
ORIGIN

Query Match 45.5%; Score 580.4; DB 12; Length 593;
Best local Similarity 99.7%; Pred. No. 2e-91;
Matches 592; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

512 CGGCATCCTCGGGACCCGGGAACCGAGCCGAGAGAGAGTGCCCGCGAAGAAAGCGGTCC 571

Db 1 CGGCATCTCGGGAGCCCGGAAACCGAGACCGAGAGAGGTGGCCCGGAGAAACGGTCC 60

QY 572 CCCAGCCGGAAGCGAGAGGGCGATGTTACGCTGCAGGGCAGGGCCGGGTGGAGAGC 631

Db 61 CCCAGCCGGAAGCGAGAGGGCGATGTTACGCTGCAGGGCAGGGCCGGGTGGAGAGC 120

QY 632 AGCCACCACCGAGAGACCGGAACAGTGTGGCGCGATGCAGTCTGAGCCTGGGAGCGAG 691

Db 121 AGCCACCACCGAGAGACCGGAACAGTGTGGCGCGATGCAGTCTGAGCCTGGGAGCGAG 180

QY 692 AGCCATGTCCTGCGCGAGAGATGGCTCAGAGACCCCGGTGATTGGATGCCCTCGAGACC 751

Db 181 AGCCATGTCCTGCGCGAGAGATGGCTCAGAGACCCCGGTGATTGGATGCCCTCGAGACC 240

QY 752 AGGCTCTCCCGGAAAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 811

Db 241 AGGCTCTCCCGGAAAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

QY 812 AGAGTACGCTACTATCACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 871

Db 301 AGAGTACGCTACTATCACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 872 GGTGTGTGAG 931

Db 361 GGTGTGTGAG 419

QY 932 TCCTACACCTCTTACAG 991

Db 420 TCCTACACCTCTTACAG 479

QY 992 GCGTGGCCAGTCAGATTTCCGACAGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1051

Db 480 GCGTGGCCAGTCAGATTTCCGACAGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539

QY 1052 GGGAGATGCAAG 1105

Db 540 GGGAGATGCAAG 593

RESULT 5
LOCUS BF471866 594 bp mRNA linear EST 04-DEC-2000
DEFINITION UT-M-BH3-awu-b-08-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
ACCESSION BF471866
VERSION BF471866.1 GI:11541049
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 594)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 92044477

COMMENT
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.
Location/Qualifiers

source
1. 594
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-awu-b-08-0-UI"
/clone_11b="NIH_BMAP_M_S4"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S1,
NIH_BMAP_M_S2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S3.2,
NIH_BMAP_M_S3.3, NIH_BMAP_M_S4. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M_S3.1, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.3 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.1, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.3
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M_S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)"

BASE COUNT 105 a 184 G 158 g 147 t
ORIGIN
Query Match 44.7%; Score 570.4; DB 12; length 594;
Best Local Similarity 99.5%; Pred. No. 1.1e-89;
Matches 593; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 555 CGCGAG 614
Db 594 CGCGAG 535
QY 615 GCGCGAG 674
Db 534 GCGCGAG 475
QY 675 TGAGCTGGAG 734
Db 474 TGAGCTGGAG 415
QY 735 GGATGCCCTCGAG 794
Db 414 GGATGCCCTCGAG 355
QY 795 TTTCCAGTTCTTAG 854
Db 354 TTTCCAGTTCTTAG 295
QY 855 GGAGAGCGCTATGT 914
Db 294 GGAGAGCGCTATGT 236
QY 915 GCGAGTGTGTGAGAAATCTTACCAACCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 974
Db 235 GCGAGTGTGTGAGAAATCTTACCAACCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 176
QY 975 GTAAAGAGACTAGATGTGCTGCCAGTCAGATTTGCCACAGTGGAGACCTTAACGCCCCC 1034
Db 175 GTAAAGAGACTAGATGTGCTGCCAGTCAGATTTGCCACAGTGGAGACCTTAACGCCCCC 116

[illegible]

RESULT 6					
BB704019					
LOCUS	BB704019	525 bp	mrna	linear	EST 11-OCT-2001
DEFINITION	BB704019	RIKEN full-length enriched, in vitro fertilized eggs Mus			
	musculus	cdna clone 7420459B08 3'			mrna sequence.

SOURCE	house mouse.
ORGANISM	Mus musculus

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (bases 1 to 525)
REFERENCE
AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,

REFERENCE

1 (bases 1 to 499)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE

9704447

COMMENT Contact: Chln, H

National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mestr@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized basal ganglia library cDNA library preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.

Seq primer: M13 Forward

POLYA=yes

FEATURES

Location/Qualifiers

1..499

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH0-akc-d-12-0-UI"

/clone_lib="NIH_BMAP_M_S1"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; The

NIH_BMAP_M_S1 library is a subtracted library derived from

a mixture of normalized libraries from ten regions of the

mouse brain (cerebellum, brain stems, olfactory bulbs,

hypothalamus, cortex, amygdala, basal ganglia, pineal

gland, striatum, hippocampus). The driver used for

subtraction consisted of a pool of 20,000 cDNA clones

obtained from non-normalized and normalized libraries of

these ten regions of the mouse brain.

TAG_LIB=NIH_BMAP_M_S1

TAG_TISSUE=basal-ganglia

TAG_SEQ=GTGAC

BASE COUNT 115 a 120 c 117 g 147 t

ORIGIN

Query Match 37.1%; Score 473.8; DB 9; Length 499;

Best Local Similarity 99.2%; Pred. No. 7.2e-73;

Matches 497; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 772 GAGCAGGACAAGAGCGCGCTGCTTCCAGTCTTAGAGCAGAAGTACGGCTACTATCAC 831
 |||||||
 Db 499 GAGCAGGACAAGAGCGCGCTGCTTCCAGTCTTAGAGCAGAAGTACGGCTACTATCAC 440
 |||||||
 QY 832 TGCAGGAGCTGCAAAATCCGCTGGAGAGCGCGCTATGTGTGTGTGTCAGGCAACAGT 891
 |||||||
 Db 439 TGCAGGAGCTGCAAAATCCGCTGGAGAGCGCGCTATGTGTGTGTGTCAGGCAACAGT 380
 |||||||
 QY 892 AAGCTGTACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCTTACACACCTTTACAGAGT 951
 |||||||
 Db 379 AAGGTG-TACTTCAAAACAGTTCTGCCGAGTGTGTGAGAAATCTTACACACCTTTACAGAGT 321
 |||||||
 QY 952 GGAGGACATCACCTGTCAAAAGTTGTAAAGAAGTATGTCCTGCCAGTCAGATTTCG 1011
 |||||||
 Db 320 GGAGGACATCACCTGTCAAAAGTTGTAAAGAAGTATGTCCTGCCAGTCAGATTTCG 261
 |||||||
 QY 1012 CCACGTGGACCCCTAAAGCCCATCGGCAAGAGCTTGTGTGGAGATGCAAGACAACG 1071
 |||||||
 Db 260 CCACGTGGACCCCTAAAGCCCATCGGCAAGAGCTTGTGTGGAGATGCAAGACAACG 201
 |||||||

QY 1072 CCTGCTCCTGGCAGACACCTTTCAGCTTCAAAATACATCATTTAGTGAGAGTCGAAACGTT 1131
 |||||||
 Db 200 CCTGCTCCTGGCAGACACCTTTCAGCTTCAAAATACATCATTTAGTGAGAGTCGAAACGTT 141
 |||||||
 QY 1132 TCTGCTAGATGGGGCTAAATGGAATGACAAGTGAAGCTTTCTCCCTCTTCACTCTTCCC 1191
 |||||||
 Db 140 TCTGCTAGATGGGGCTAAATGGAATGACAAGTGAAGCTTTCTCCCTCTTCACTCTTCCC 81
 |||||||
 QY 1192 TTTCCAAATCTTCAATGACAGACAGTGTACTTGGATATAAAGCCTGTGAATAAAGGTA 1251
 |||||||
 Db 80 TTTCCAAATCTTCAATGACAGACAGTGTACTTGGATATAAAGCCTGTGAATAAAGGTA 22
 |||||||
 QY 1252 TTGCAACAACAAAAA 1272
 |||||||
 Db 21 TTGCAACAACAAAAA 1

RESULT 8

BB703259 491 bp mRNA linear EST 11-OCT-2001

LOCUS

BB703259 RIKEN full-length enriched, in vitro fertilized eggs Mus

DEFINITION

musculus cDNA clone 7420449J15 3', mRNA sequence.

ACCESSION

BB703259

VERSION

BB703259.1 GI:16052094

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AUTHORS

1 (bases 1 to 491)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shitaki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takeku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watabiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)

TITLE

Unpublished (2001)

JOURNAL

Contact: Yoshihide Hayashizaki

COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic

JOURNAL

Sciences Center(GSC), Yokohama Institute

COMMENT

The Institute of Physical and Chemical Research (RIKEN)

JOURNAL

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

COMMENT

Tel: 81-45-503-9222

JOURNAL

Fax: 81-45-503-9216

COMMENT

Email: genome-res@gsr.riken.go.jp,

JOURNAL

URL: http://genome.gsc.riken.go.jp/

COMMENT

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

JOURNAL

Normalization and subtraction of cap-trapper-selected cDNAs to

COMMENT

prepare full-length cDNA libraries for rapid discovery of new

JOURNAL

genes. Genome Res. 10 (10), 1617-1630 (2000)

COMMENT

wagl,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuyama,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

JOURNAL

RIKEN integrated sequence analysis (RISA) system--384-format

COMMENT

sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.

COMMENT

Computer-based methods for the mouse full-length cDNA

JOURNAL

encyclopedia: real-time sequence clustering for construction of a

COMMENT

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

JOURNAL

Please visit our web site (http://genome.gsc.riken.go.jp) for

COMMENT

further details.

JOURNAL

e mouse tissues.

COMMENT

Location/Qualifiers

JOURNAL

1..491

COMMENT

source

FEATURES

Location/Qualifiers

COMMENT

1..491

JOURNAL

source

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420449J15"
/clone_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/note="Site_1: Sall; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATCTCGAGTTATTAATTAATCCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified Bluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sall; 3' end: BamHI"

BASE COUNT 138 a 118 c 119 g 116 t

ORIGIN

Query Match 35.1%; Score 448.4; DB 10; Length 491;
Best Local Similarity 98.2%; Pred. No. 1.9e-68;
Matches 485; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 763 CAAGCAGGAGCAGCAGCAGGAGCGCCCTGCTTCCAGTCTTAGAGCAGAGTACGGC 822
DB 1 CAAGCAGGAGCAGCAGCAGGAGCGCCCTGCTTCCAGTCTTAGAGCAGAGTACGGC 60
QY 823 TACTATCACTGCAAGACTGCAAAATCCGGTGGAGAGCGCCCTATGTGTGTGTGAG 882
DB 61 TACTATCACTGCAA-GACTGCAAAATCCGGTGGAGAGCGCCCTATGTGTGTGTGAG 119
QY 883 GGCACCACTGAGGTGTTACTTCAACAGTTCGCCGAGTGTGAGAAATCCTACAAACC 942
DB 120 GGCACCACTGAGGTG-TACTTAAACAGTTCGCCGAGTGTGAGAAATCCTACAAACC 178
QY 943 TTACAGAGTGAGGACATCACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCCAGT 1002
DB 179 TTACAGAGTGAGGACATCACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCCAGT 238
QY 1003 CAGATTTCGCCACGTGGAACCCCTAAACGCCCCCAATCCGCAAGACTTGTGGGAGATGCAA 1062
DB 239 CAGACTTCGCCACGTGGAACCCCTAAACGCCCCCAATCCGCAAGACTTGTGGGAGATGCAA 298
QY 1063 GGACAAACGGCTGTCTCGACAGCAGCAGCTTCAAGTTCAATATCATTTAGTGAGAGTC 1122
DB 299 GGACAAACGGCTGTCTCGACAGCAGCAGCTTCAAGTTCAATATCATTTAGTGAGAGTC 358
QY 1123 GAAACGTTTCTGCTAGATGGGGCTAATGGAATGGACAAGTGAAGCTTCTCCCTCTTCA 1182
DB 359 GAAACGTTTCTGCTAGATGGGGCTAATGGAATGGACAAGTGAAGCTTCTCCCTCTTCA 418
QY 1183 CCTCTTCCCTTTCCAAATTTCTATGACAGACAGTGTACTTGATATAAAGCCTGTGAA 1242
DB 419 CCTCTTCCCTTTCCAAATTTCTATGACAGACAGTGT-TACTTGATATAAAGCCTGTGAA 477
QY 1243 TAAAGGTATTGCA 1256
DB 478 TAAAGGTATTGAA 491

RESULT 9
BB704648 BB704648 491 bp mRNA linear EST 11-OCT-2001
LOCUS

DEFINITION BB704648 RIKEN full-length enriched, in vitro fertilized eggs Mus
musculus cDNA clone 7420466L07 3', mRNA sequence.
ACCESSION BB704648
VERSION BB704648.1 GI:16053483
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 491)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

TITLE Unpublished (2001)

JOURNAL Contact: Yoshinori Hayashizaki
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscl.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
waji,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES

source

Location/Qualifiers
1..491
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420466L07"
/clone_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/note="Site_1: Sall; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the

RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT	135 a	133 c	156 g	149 t	17 others
ORIGIN					

Query Match	31.0%;	Score 396.4;	DB 9;	Length 590;
Best Local Similarity	86.1%;	Pred. No. 1.9e-59;		
Matches 464;	Conservative 0;	Mismatches 71;	Indels 4;	Gaps 3;

QY	713	TGGCTCAGGACCCCGGTGATTCGGATGCCCCCTCGAGACCAAGGCCCTCCCCGCCAAAGCACGG	772
Db	542	TGGNCTNAGACCCCCCGTTATTCGATGCCCCCTGGAACACAGGCNTCCNCNCAANNCCGGACC	483
QY	773	AGCAGGACAAGAGCGCCCTGCGTTTCCAGTCTTAGAGCAGAGTACGG - CTACTATCA	830
Db	482	AGNACAAGAGCGGCCCTGCNTTTTCCAGCTCTTAGAGCACAAAGCTACGNCCCTACTATCA	423
QY	831	CTGCAAGGACTGCAAAATCCGCTGGGAGAGCGCCTATGTG - TGGTGTGTGACGGGCACCA	889
Db	422	CTGCAAGGACTGCAAAATCCGNTGGGAGANCNCCTATGTGTGTGTGTGACGGCCACC	363
QY	890	GTAAGGTGTACTTCAAAACAGTTCTGCCGAGTGTGAGAAATCCTACAACCCCTTACAGA	949
Db	362	AGTAAGGTGTACTTCAAAACAGTTCTGCCGAGTGTGAGAAATCCTACAACCNCTTACAGA	303
QY	950	GTCGAGGACATCACCTGTCAAAAGTTGTAAAGAACTAGATGTGCTGCCAGTCAGATT	1009
Db	302	GTCGAGGACATCACCTGTCAAAAGTTGTAAAGAACTAGATGTGCTGCCAGTCAGACTT	243
QY	1010	CGCCACGTGGACCCTAAACGCCCCCATCGGCAAGACTTGTGTGGAGATGCAAGGACAAA	1069
Db	242	CGCCACGTGGACCCTAAACGCCCCCATCGGCAAGACTTGTGTGGAGATGCAAGGACAAA	183
QY	1070	CGCCTGTCTCGCAGACGACCTTCAGCTTCAATACATCATTTAGTGAGAGTCGAAAACG	1129
Db	182	CGCCTGTCTCGCAGACGACCTTNAAGCTTCAATACATCATTTAGTGAGAGTCGAAAACG	123
QY	1130	TTTCTGCTAGATGGGGCTAATGGAATGGACAAGTGAGCTTTCTCCCCCTCTTACCCTTTC	1189
Db	122	TTTCTGNTAGATGGGGCTAATGGAATGGACAAGTGAGCTTTCTCCCCCTCTTACCCTTTC	63
QY	1190	CCTTTCCAATTTCTTCATGACAGACAGTGTACTTGGATATTAAGCCTGTGAATAAAAG	1248
Db	62	CCTTTCCAATTTCTTCATGACAGACAGTGT - TACTTGGATATTAAGCCTGTGAATAAAAG	5

RESULT 13	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
BE946858	BE946858	450 bp mRNA linear EST 03-OCT-2000	UI-M-BH3-awu-b-08-0-UI.s1	NIH_BMAP_M_S4	Mus musculus cDNA clone	UI-M-BH3-awu-b-08-0-UI 3', mRNA sequence.							
BE946858	BE946858		BE946858	GI:10524617	EST.	house mouse.	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Bonaldo,M.F., Lennon,G. and Soares,M.B.	Normalization and subtraction: two approaches to facilitate gene discovery	Genome Res. 6 (9), 791-806 (1996)	97044477	Contact: Chin, H

is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements. Seq primer: M13 Forward POLYA=NO.

FEATURES	Location/Qualifiers
source	1. .450

BASE COUNT	62 a	168 c	156 g	63 t	1 others
ORIGIN					
Query Match		30.8%;	Score 393;	DB 12;	Length 450;
Best Local Similarity		99.7%;	Pred. No. 8.2e-59;		
Matches 393;	Conservative	0;	Mismatches	1;	Indels 0;
					Gaps 0;
QY	3	GGCGGGCGAGGCGCGGAGCGCACCCATGTTCCCGGCGAGCAGCTTCCACCCCTGCCCCGA	62		
Db	57	GGCGGGCGAGGCGCGGAGCGCACCCATGTTCCCGGCGAGCAGCTTCCACCCCTGCCCCGA	116		
QY	63	TCCTTATCCGACAGGCCACCAAGCCGGGATGGCTGAGGTTGGAGCCAGGGGCTGCCG	122		
Db	117	TCCTTATCCGACAGGCCACCAAGCCGGGATGGCTGAGGTTGGAGCCAGGGGCTGCCG	176		
QY	123	ACCGCGCCCCCTCTCTTCTTCCCGGCTACAGACAGCTCATGGCCGGGAGTACGTGA	182		
Db	177	ACCGCGCCCCCTCTCTTCTTCCCGGCTACAGACAGCTCATGGCCGGGAGTACGTGA	236		
QY	183	CAGCCACCAGCGGGCACAGCTCATGGCCCTGCTGTCCCGCATGGGTTCCCGGTCGTCAG	242		
Db	237	CAGCCACCAGCGGGCACAGCTCATGGCCCTGCTGTCCCGCATGGGTTCCCGGTCGTCAG	296		
QY	243	CAGCCGTGACGCTGCGGTGCAGGTTGAACCCGCGCGCGCAGCCCTCGGTGCAGTGTTCAC	302		
Db	297	CAGCCGTGACGCTGCGGTGCAGGTTGAACCCGCGCGCGCAGCCCTCGGTGCAGTGTTCAC	356		
QY	303	CGGCGCGCACGCTGCAGCCTGCAGGGTGCAGCCAGCCCGGACGCGCGATCGGGTTC	362		

Db 357 CGGGCGCCGACAGCTGCAGCCTGCAGGTCGCCAGCCAGCCCGACGCCCGCATCGGCTTC 416

QY 363 CTGTCAACCCCGTGGCCAGCGCGCGCGGAGA 396
|||||
Db 417 CTGTCAACCCCGTGGNCACGCCGCGCGGAGA 450

RESULT 14
BB704449 419 bp mRNA linear EST 11-OCT-2001
LOCUS BB704449 RIKEN full-length enriched, in vitro fertilized eggs Mus
DEFINITION musculus cDNA clone 7420464A04 3', mRNA sequence.
ACCESSION BB704449
VERSION BB704449.1 GI:16053284
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 419)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

JOURNAL Unpublished (2001)
COMMENT Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscl.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10
(11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES
Source location/Qualifiers
1. 419
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420464A04"
/clone_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCGAGTTAATTAAATTAATCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from lambda
FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

BASE COUNT 115 a 100 c 100 g 104 t

ORIGIN

Query Match 30.7%; Score 392.2; DB 10; Length 419;
Best Local Similarity 98.8%; Pred. No. 1.2e-58;
Matches 416; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 834 CAAGACTGCAAAATCCGGTGGAGAGCGCTATGTGTGTGTGTGCAGGGCACCAGTAA 893
|||||
Db 1 CAAGACTGCAAAATCCGGTGGAGAGCGCTATGTGTGTGTGTGCAGGGCACCAGTAA 60

QY 894 GGTGTACTTCAACAGTTCCTGCCGAGTGTGTGAGAAATCCTACACCCCTTACAGAGTGG 953
|||
Db 61 GGT-TTACTTCAACAGTTCCTGCCGAGTGTGTGAGAAATCCTACACCCCTTACAGAGTGG 119

QY 954 AGGACATCACTGTCAAAAGTTGTAAGAAGACTAGATGTGCTGCCAGTCAGATTTCGCC 1013
|||||
Db 120 AGGACATCACTGTCAAAAGTTGTAAGAAGACTAGATGTGCTGCCAGTCAGATTTCGCC 179

QY 1014 ACGTGGACCTAAACGCCCCCATCGGCAAGACTTGTGTGGAGATGCAAGGACAAACGCC 1073
|||||
Db 180 ACGTGGACCTAAACGCCCCCATCGGCAAGACTTGTGTGGAGATGCAAGGACAAACGCC 239

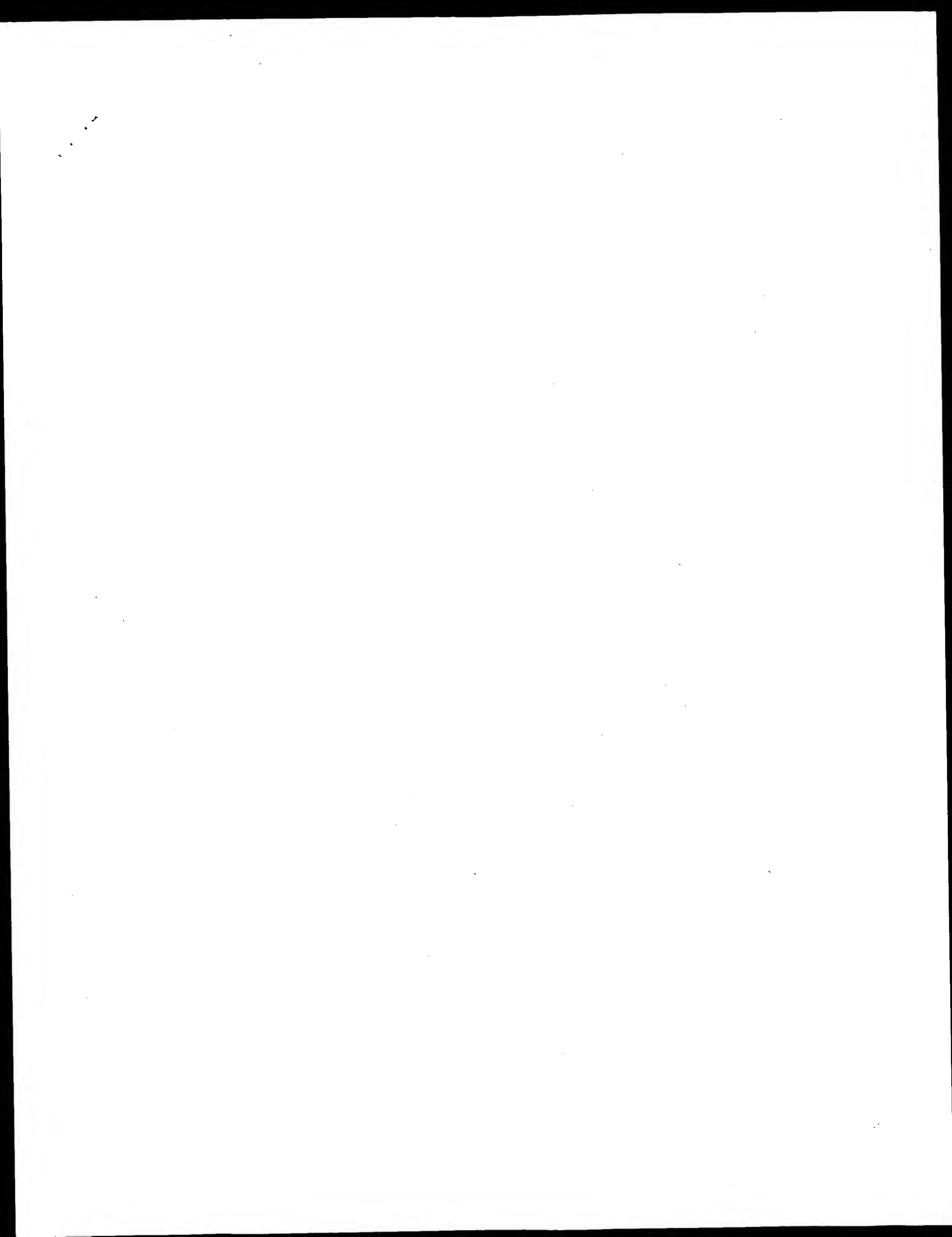
QY 1074 TGTCTCGCAGACAGCAGCTTCAGCTTCAAAATACATCATTTAGTGAAGTGGAAACGTTTC 1133
|||||
Db 240 TGTCTCGCAGACAGCAGCTTCAGCTTCAAAATACATCATTTAGTGAAGTGGAAACGTTTC 299

QY 1134 TGCTAGATGGGGCTTAATGGAATGACAAAGTGAGCTTTCTCCCTCTTCACTCTTCCTT 1193
|||||
Db 300 TGCTAGATGGGGCTTAATGGAATGACAAAGTGAGCTTTCTCCCTCTTCACTCTTCCTT 359

QY 1194 TCCAAATTTCTCATGACAGACAGAGTGTACTTGATATAAAGCCTGTGAATAAAGGTATT 1253
|||||
Db 360 TCCAAATTTCTCATGACAGACAGAGTGTACTTGATATAAAGCCTGTGAATAAAGGTATT 418

QY 1254 G 1254
|
Db 419 G 419

RESULT 15
BB703869 521 bp mRNA linear EST 11-OCT-2001
LOCUS BB703869 RIKEN full-length enriched, in vitro fertilized eggs Mus
DEFINITION musculus cDNA clone 7420457C21 3', mRNA sequence.
ACCESSION BB703869
VERSION BB703869.1 GI:16052704
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 521)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2003, 20:01:55 ; Search time 3366 Seconds

(without alignments)
11041.082 Million cell updates/sec

Title: US-09-844-864-1

Perfect score: 1277
Sequence: 1 aaagcggcgcgagcgcgga.....acaaaaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database :

GenEmbl: *	1: gb_ba: *
2: gb_htg: *	2: gb_htg: *
3: gb_in: *	3: gb_in: *
4: gb_om: *	4: gb_om: *
5: gb_ov: *	5: gb_ov: *
6: gb_pat: *	6: gb_pat: *
7: gb_ph: *	7: gb_ph: *
8: gb_pl: *	8: gb_pl: *
9: gb_pr: *	9: gb_pr: *
10: gb_ro: *	10: gb_ro: *
11: gb_sts: *	11: gb_sts: *
12: gb_sy: *	12: gb_sy: *
13: gb_un: *	13: gb_un: *
14: gb_vl: *	14: gb_vl: *
15: em_ba: *	15: em_ba: *
16: em_fun: *	16: em_fun: *
17: em_hum: *	17: em_hum: *
18: em_in: *	18: em_in: *
19: em_mu: *	19: em_mu: *
20: em_om: *	20: em_om: *
21: em_or: *	21: em_or: *
22: em_ov: *	22: em_ov: *
23: em_pat: *	23: em_pat: *
24: em_ph: *	24: em_ph: *
25: em_pl: *	25: em_pl: *
26: em_ro: *	26: em_ro: *
27: em_sts: *	27: em_sts: *
28: em_un: *	28: em_un: *
29: em_vl: *	29: em_vl: *
30: em_htg_hum: *	30: em_htg_hum: *
31: em_htg_in: *	31: em_htg_in: *
32: em_htg_other: *	32: em_htg_other: *
33: em_htg_mus: *	33: em_htg_mus: *
34: em_htg_pln: *	34: em_htg_pln: *
35: em_htg_rnd: *	35: em_htg_rnd: *
36: em_htg_mam: *	36: em_htg_mam: *
37: em_htg_vrt: *	37: em_htg_vrt: *
38: em_sy: *	38: em_sy: *
39: em_htgo_hum: *	39: em_htgo_hum: *
40: em_htgo_mus: *	40: em_htgo_mus: *
41: em_htgo_other: *	41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	799.4	62.6	200535	10	AL671880	AL671880 Mouse DNA
2	689.8	54.0	144963	2	AC107686	AC107686 Mus muscu
3	657.4	51.5	67535	2	AC122733	AC122733 Mus muscu
4	615.4	48.2	173007	2	AC125993	AC125993 Rattus no
5	585.2	45.8	89551	2	AC127083	AC127083 Rattus no
6	452.4	35.4	169908	2	AC126519	AC126519 Rattus no
7	360	28.2	212848	2	AC108848	AC108848 Mus muscu
8	299.8	23.5	212848	2	AC108848	AC108848 Mus muscu
9	194.8	15.3	89551	2	AC127083	AC127083 Rattus no
10	147.6	11.6	138890	9	AC007970	AC007970 Homo sapi
11	147.6	11.6	151961	2	AC037443	AC037443 Homo sapi
12	147.6	11.6	204268	2	AC092163	AC092163 Homo sapi
13	134.6	10.5	187578	9	AC096952	AC096952 Homo sapi
14	73.2	5.7	694	9	HSA335296	AJ335296 Homo sapi
15	72.2	5.7	125020	9	AF429315	AF429315 Homo sapi
16	66.8	5.2	146431	2	AC109670	AC109670 Rattus no
17	64.4	5.0	172433	2	AC118973	AC118973 Rattus no
18	63.2	4.9	125020	9	AF429315	AF429315 Homo sapi
19	62.8	4.9	194873	2	AC125563	AC125563 Rattus no
20	62.6	4.9	1393	11	PM11H126	AL684264 Penicillium
21	62.	4.9	136551	2	AC048354	AC048354 Homo sapi
22	62	4.9	138467	2	AC111697	AC111697 Rattus no
23	59	4.6	185263	2	AC129765	AC129765 Rattus no
24	58.6	4.6	194936	2	AC115880	AC115880 Mus muscu
25	58.4	4.6	36394	1	SCD35	AL160312 Streptomy
26	58	4.5	92310	2	AC108380	AC108380 Pan trogl
27	58	4.5	97065	2	AC121735	AC121735 Rattus no
28	57.2	4.5	62847	2	AC125560	AC125560 Rattus no
29	57.2	4.5	83629	2	AC110485	AC110485 Rattus no
30	57.2	4.5	144053	2	AC092857	AC092857 Rattus no
31	57.2	4.5	150090	2	AC121750	AC121750 Rattus no
32	57.2	4.5	150238	2	CNS08CA9	AL831796 Oryza sat
33	57	4.5	92055	2	AC105692	AC105692 Rattus no
34	57	4.5	115553	2	AC126077	AC126077 Rattus no
35	56.8	4.4	73041	2	AC127648	AC127648 Rattus no
36	56.8	4.4	146431	2	AC109670	AC109670 Rattus no
37	56.6	4.4	157448	2	AC131139	AC131139 Rattus no
38	56.4	4.4	136788	2	AC117044	AC117044 Rattus no
39	56.2	4.4	159677	2	AC113802	AC113802 Rattus no
40	56.2	4.4	184402	2	AC127041	AC127041 Rattus no
41	55.8	4.4	179608	2	AC128497	AC128497 Rattus no
42	55.6	4.4	991	11	PM12H12B	AL684455 Penicillium
43	55.6	4.4	134366	2	AC119661	AC119661 Rattus no
44	55.6	4.4	136014	2	AC126531	AC126531 Rattus no
45	55.6	4.4	151349	10	AL355176	AL355176 Mouse DNA

ALIGNMENTS

RESULT 1
AL671880
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AL671880 20535 bp DNA linear ROD 03-AUG-2002
 Mouse DNA sequence from clone RP23-384C22 on chromosome X, complete
 sequence.
 AL671880
 AL671880.14 GI:22204349

SOURCE ORGANISM	house mouse. Mus musculus
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

REFERENCE

TITLE Direct Submission

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Landers, T., Lehoczkzy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riebeck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainotti, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

COMMENT

Submitted: 21-APR-2002 Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 21, 2002 this sequence version replaced gi:18308388.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L17749

Center clone name: 228_B_12

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 141008 bases at least Q40

Consensus quality: 142434 bases at least Q30

Consensus quality: 142970 bases at least Q20

Insert size: 14000; agarose-1p

Insert size: 143263; sum-of-contigs

Quality coverage: 8.2 in Q20 bases; agarose-1p

Quality coverage: 8.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 630: contig of 630 bp in length

* 631 730: gap of 100 bp

* 731 2440: contig of 1710 bp in length

* 2441 2540: gap of 100 bp

* 2541 4613: contig of 2073 bp in length

* 4614 4713: gap of 100 bp

* 4714 6263: contig of 1550 bp in length

* 6264 6363: gap of 100 bp

* 6364 9200: contig of 2837 bp in length

* 9201 9300: gap of 100 bp

* 9301 12704: contig of 3404 bp in length

* 12705 12804: gap of 100 bp

* 12805 15990: contig of 3186 bp in length

* 15991 16090: gap of 100 bp

* 16091 20096: contig of 4006 bp in length

FEATURES

source

* 20097 20196: gap of 100 bp
* 20197 23757: contig of 3561 bp in length
* 23758 23857: gap of 100 bp
* 23858 35345: contig of 11488 bp in length
* 35346 35445: gap of 100 bp
* 35446 43937: contig of 8492 bp in length
* 43938 44037: gap of 100 bp
* 44038 54685: contig of 10648 bp in length
* 54686 54785: gap of 100 bp
* 54786 64286: contig of 9501 bp in length
* 64287 64386: gap of 100 bp
* 64387 77535: contig of 13149 bp in length
* 77536 77635: gap of 100 bp
* 77636 91482: contig of 13847 bp in length
* 91483 91582: gap of 100 bp
* 91583 103390: contig of 11808 bp in length
* 103391 103490: gap of 100 bp
* 103491 124554: contig of 21064 bp in length
* 124555 124655: gap of 100 bp
* 124655 144963: contig of 20309 bp in length.

Location/Qualifiers

1. 144963

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="RP24-228B12"

/clone_11b="RPCI-24 Male Mouse BAC"

1. 630

/note="assembly-fragment"

731. 2440

/note="assembly-fragment"

2541. 4613

/note="assembly-fragment"

4714. 6263

/note="assembly-fragment"

6364. 9200

/note="assembly-fragment"

9301. 12704

/note="assembly-fragment"

12805. 15990

/note="assembly-fragment"

16091. 20096

/note="assembly-fragment"

20197. 23757

/note="assembly-fragment"

23858. 35345

/note="assembly-fragment"

35446. 43937

/note="assembly-fragment"

44038. 54685

/note="assembly-fragment"

54786. 64286

/note="assembly-fragment"

64387. 77535

/note="assembly-fragment"

77636. 91482

/note="assembly-fragment"

91583. 103390

/note="assembly-fragment"

103491. 124554

/note="assembly-fragment"

124655. 144963

/note="assembly-fragment"

BASE COUNT 39417 a 30598 c 30931 g 42311 t 1706 others

Query Match 54.0%; Score 689.8; DB 2; Length 144963;
Best Local Similarity 95.0%; Pred. No. 5.9e-132;
Matches 761; Conservative 0; Mismatches 22; Indels 18; Gaps 4;

QY 3 GCGGGCGAGCGCGGAGCAGCACCATTGTTCCCGGAGACAGTTCACACCCCTGCCGCA 62
DB 142184 GCGGGCGAGCGCGGAGCAGCACCATTGTTCCCGGAGACAGTTCACACCCCTGCCGCA 142243

QY	63	TCCTTATCCGACAGGCCACCAAAAGCCGGGATGGCTGGAGGTTCCGAGCCAGGGGCTGCCG	122
Db	142244	TCCTTATCCGACAGGCCACCAAAAGCCGGGATGGCTGGAGGTTCCGAGCCAGGGGCTGCCG	142303
QY	123	ACCCGGCCCCCTCTCTTCTCTCCCGGCTACAGACAGCTCATGGCCCGGAGTACGTCGA	182
Db	142304	ACCCGGCCCCCTCTCTTCTCTCCCGGCTACAGACAGCTCATGGCCCGGAGTACGTCGA	142363
QY	183	CAGCCACCAGCGGGCACAGCTCATGGCCCTGCTGTGCGGATGGGTCCCGGTGGTCAG	242
Db	142364	CAGCCACCAGCGGGCACAGCTCATGGCCCTGCTGTGCGGATGGGTCCCGGTGGTCAG	142423
QY	243	CAGCCGTGACGCTGCGGTGCAGGTGAACCCGCGCCGACGCGCTCGGTGCAGTGTTCCT	302
Db	142424	CAGCCGTGACGCTGCGGTGCAGGTGAACCCGCGCCGACGCGCTCGGTGCAGTGTTCCT	142483
QY	303	CGGGCGCCGACAGCTGCAGCCTGACGGGTGCCAGCCAGCCCGACGCGCGATCGGGTTC	362
Db	142484	CGGGCGCCGACAGCTGCAGCCTGACGGGTGCCAGCCAGCCCGACGCGCGATCGGGTTC	142543
QY	363	CTGTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCGCGATCCTTGACAGCCGTAGC	422
Db	142544	CTGTCAACCCCGTGGCCACGCGCGCGGAGATCCCGCGATCCTTGACAGCCGTAGC	142603
QY	423	CCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCATGTGAGGTTGCGGAGCAG	482
Db	142604	CCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCTCTCTCATGTGAGGTTGCGGAGCAG	142658
QY	483	GCAGACACCCACGAAGGGAGAGGGAGCCCGGCATCCTCGGGGGAACCCGGAAACCGAGCC	542
Db	142659	GCAGACACCCACGAAGGGAGAGGGAGCCCGGCATCCTCGGGGGAACCCGGAAACCGAGCC	142718
QY	543	GAGAGAGGTGGCCGCGAGAAAGCGGTCCCCAGCCGCGGAAGCAGAGAGGGCGATGTCA	602
Db	142719	GAGAGAGGTGGCCGCGAGAAAGCGGTCCCCAGCCGCGGAAGCAGAGAGGGCGATGTCA	142778
QY	603	GGCTGCAGGGGACGCGGGTGGGAGCAGCAGCCACCAACCGGAGAACCGGAACAGTGTGC	662
Db	142779	GGCTGCAGGGGACGCGGGTGGGAGCAGCAGCCACCAACCGGAGAACCGGAACAGTGTGC	142838
QY	663	GGCGATGCAGTCTGAGCCTGGGAGCGAGAGCCATGTCTTCCCGCAGAGATGGCTCAGGA	722
Db	142839	GGTGTATGCAGTCTGAGCCTGGGAGTGAAGAGCTGTGTCTCTGCACAGAGATGGCTCAGGA	142898
QY	723	CCCCGGTGATTGGAATGCCCTTCGAGACCAAGGCTCCCGCGAAAGCACGGAGCAGACAA	782
Db	142899	CCCCGGTGACTCGGATGCC-----CCTCCCGCAAAAGCACCAAGCAGAGACAA	142945
QY	783	GGAGCGCCTGCGTTTCCAGTT 803	
Db	142946	GGAGCTCTCGCTTTCAGGT 142966	

RESULT	3
ACI122733	
LOCUS	67535 bp DNA linear HTG 25-MAY-2002
DEFINITION	Mus musculus clone RP24-506B15, LOW-PASS SEQUENCE SAMPLING.
ACCESSION	ACI122733
VERSION	ACI122733.1 GI:21206317
KEYWORDS	HTG; HTGS_PHASE0.
SOURCE	Mus musculus.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 67535)
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE	Mus musculus, clone RP24-506B15
JOURNAL	Unpublished
REFERENCE,	2 (bases 1 to 67535)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

TITLE
JOURNAL
COMMENT

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
Lamazares, R., Landers, T., Lehoczký, J., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N.,
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (25-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17506
Center clone name: 506_B_15

* NOTE: This record contains 84 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allows overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

*	1	669: contig of 669 bp in length
*	670	769: gap of 100 bp
*	770	1486: contig of 717 bp in length
*	1487	1586: gap of 100 bp
*	1587	2304: contig of 718 bp in length
*	2305	2404: gap of 100 bp
*	2405	3104: contig of 700 bp in length
*	3105	3204: gap of 100 bp
*	3205	3870: contig of 666 bp in length
*	3871	3970: gap of 100 bp
*	3971	4672: contig of 702 bp in length
*	4673	4772: gap of 100 bp
*	4773	5450: contig of 678 bp in length
*	5451	5550: gap of 100 bp
*	5551	6248: contig of 698 bp in length
*	6249	6348: gap of 100 bp
*	6349	7053: contig of 705 bp in length
*	7054	7153: gap of 100 bp
*	7154	7865: contig of 712 bp in length
*	7866	7965: gap of 100 bp
*	7966	8668: contig of 703 bp in length
*	8669	8768: gap of 100 bp
*	8769	9479: contig of 711 bp in length
*	9480	9579: gap of 100 bp
*	9580	10301: contig of 722 bp in length
*	10302	10401: gap of 100 bp
*	10402	11115: contig of 714 bp in length
*	11116	11215: gap of 100 bp
*	11216	11878: contig of 663 bp in length

* 11879 11978: gap of 100 bp
* 11979 12686: contig of 708 bp in length
* 12687 12786: gap of 100 bp
* 12787 13495: contig of 709 bp in length
* 13496 13595: gap of 100 bp
* 13596 14306: contig of 711 bp in length
* 14307 14406: gap of 100 bp
* 14407 15103: contig of 697 bp in length
* 15104 15203: gap of 100 bp
* 15204 15921: contig of 718 bp in length
* 15922 16021: gap of 100 bp
* 16022 16696: contig of 675 bp in length
* 16697 16796: gap of 100 bp
* 16797 17519: contig of 723 bp in length
* 17520 17619: gap of 100 bp
* 17620 18337: contig of 718 bp in length
* 18338 18437: gap of 100 bp
* 18438 19132: contig of 695 bp in length
* 19133 19232: gap of 100 bp
* 19233 19906: contig of 674 bp in length
* 19907 20006: gap of 100 bp
* 20007 20705: contig of 699 bp in length
* 20706 20805: gap of 100 bp
* 20806 21498: contig of 693 bp in length
* 21499 21598: gap of 100 bp
* 21599 22302: contig of 704 bp in length
* 22303 22402: gap of 100 bp
* 22403 23108: contig of 706 bp in length
* 23109 23208: gap of 100 bp
* 23209 23927: contig of 719 bp in length
* 23928 24027: gap of 100 bp
* 24028 24748: contig of 721 bp in length
* 24749 24848: gap of 100 bp
* 24849 25566: contig of 718 bp in length
* 25567 25666: gap of 100 bp
* 25667 26370: contig of 704 bp in length
* 26371 26470: gap of 100 bp
* 26471 27169: contig of 699 bp in length
* 27170 27269: gap of 100 bp
* 27270 27963: contig of 694 bp in length
* 27964 28063: gap of 100 bp
* 28064 28757: contig of 694 bp in length
* 28758 28857: gap of 100 bp
* 28858 29570: contig of 713 bp in length
* 29571 29670: gap of 100 bp
* 29671 30383: contig of 713 bp in length
* 30384 30483: gap of 100 bp
* 30484 31189: contig of 706 bp in length
* 31190 31289: gap of 100 bp
* 31290 32018: contig of 729 bp in length
* 32019 32118: gap of 100 bp
* 32119 32843: contig of 725 bp in length
* 32844 32943: gap of 100 bp
* 32944 33614: contig of 671 bp in length
* 33615 33714: gap of 100 bp
* 33715 34436: contig of 722 bp in length
* 34437 34536: gap of 100 bp
* 34537 35257: contig of 721 bp in length
* 35258 35357: gap of 100 bp
* 35358 36063: contig of 706 bp in length
* 36064 36163: gap of 100 bp
* 36164 36856: contig of 693 bp in length
* 36857 36956: gap of 100 bp
* 36957 37661: contig of 705 bp in length
* 37662 37761: gap of 100 bp
* 37762 38464: contig of 703 bp in length
* 38465 38564: gap of 100 bp
* 38565 39269: contig of 705 bp in length
* 39270 39369: gap of 100 bp
* 39370 40052: contig of 683 bp in length
* 40053 40152: gap of 100 bp
* 40153 40863: contig of 711 bp in length
* 40864 40963: gap of 100 bp

* 40964 41682: contig of 719 bp in length
* 41683 41782: gap of 100 bp
* 41783 42490: contig of 708 bp in length
* 42491 42590: gap of 100 bp
* 42591 43311: contig of 721 bp in length
* 43312 43411: gap of 100 bp
* 43412 44137: contig of 726 bp in length
* 44138 44237: gap of 100 bp
* 44238 44946: contig of 709 bp in length
* 44947 45046: gap of 100 bp
* 45047 45758: contig of 712 bp in length
* 45759 45858: gap of 100 bp
* 45859 46553: contig of 695 bp in length
* 46554 46653: gap of 100 bp
* 46654 47366: contig of 713 bp in length
* 47367 47466: gap of 100 bp
* 47467 48179: contig of 713 bp in length
* 48180 48279: gap of 100 bp
* 48280 48944: contig of 665 bp in length
* 48945 49044: gap of 100 bp
* 49045 49763: contig of 719 bp in length
* 49764 49863: gap of 100 bp
* 49864 50570: contig of 707 bp in length
* 50571 50670: gap of 100 bp
* 50671 51392: contig of 722 bp in length
* 51393 51492: gap of 100 bp
* 51493 52184: contig of 692 bp in length
* 52185 52284: gap of 100 bp
* 52285 52992: contig of 708 bp in length
* 52993 53092: gap of 100 bp
* 53093 53787: contig of 695 bp in length
* 53788 53887: gap of 100 bp
* 53888 54595: contig of 708 bp in length
* 54596 54695: gap of 100 bp
* 54696 55404: contig of 709 bp in length
* 55405 55504: gap of 100 bp

Query Match 51.5%; Score 657.4; DB 2; Length 67535;
Best Local Similarity 86.8%; Pred. No. 2.8e-125;
Matches 669; Conservative 0; Mismatches 101; Indels 1; Gaps 1;

QY 29 TGTTCGGGGGAGCAGCTTCCACCCCTGCGCATCTTATCCGAGCCACCAAGCCG 88
Db 18464 TGTTCGGGGGAGCAGCTTCCACCCCTGCGCATCTTATCCGAGCCACCAAGCCG 18523
QY 89 GGGATGGCTGAGGTTGCGAGCCAGGGGTCGCGACCCGCGCCCTCTCTCCCGG 148
Db 18524 GGGATGGCTGAGGTTGCGAGCCAGGGGTCGCGACCCGCGCCCTCTCTCCCGG 18583
QY 149 GCTACAGACAGCTCATGCGCGGAGTACGTGACAGCCAGCCAGCGGGACAGCTCATGG 208
Db 18584 GCTACAGACAGCTCATGCGCGGAGTACGTGACAGCCAGCCAGCGGGACAGCTCATGG 18643
QY 209 CCTGCTGTGCGGATGGGTCCCGGTGCGTACAGACCCGTGACGCTGCGGTGAGTGA 268
Db 18644 CCTGCTGTGCGGATGGGTCCCGGTGCGTACAGACCCGTGACGCTGCGGTGAGTGA 18703
QY 269 ACCCGCGCGGAGCGCTCGGTGACAGTTCACCTGCGGCGCCGACAGCTGACGCTGCGAG 328
Db 18704 ACCCGCGCGGAGCGCTCGGTGACAGTTCACCTGCGGCGCCGACAGCTGACGCTGCGAG 18763
QY 329 GGTGCGGAGCAGCCCGGAGCGCCGATCGGGTCTCTGTCAACCCCGTGGCCAGCGCGCG 388
Db 18764 GGTGCGGAGCAGCCCGGAGCGCCGATCGGGTCTCTGTCAACCCCGTGGCCAGCGCGCG 18823
QY 389 CCGGAGATCCCGCGGATCTCTGCGACAGCTAGCCCGCTTCTGCTGACCTTCTGTG 448
Db 18824 CCGGAGATCCCGCGGATCTCTGCGACAGCTAGCCCGCTTCTGCTGACCTTCTGTG 18883
QY 449 GCCTCTCTCTCTACTGAGGTTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 508
Db 18884 GCCTCTCTCTCTACTGAGGTTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18943

QY 509 GCCCGCATCTCCGGGAGACCCGGAGACCCGAGAGAGTGGCCGCGAGAAAGCGG 568
|||||
Db 18944 GCCCGCATCTCCGGGAGACCCGGAGACCCGAGAGAGTGGCCGCGAGAAAGCGG 19003
QY 569 TCCCCAGCCGCGAAGCGAGAGGGCGATGTTTCAGGCTGCAGGGCA-GCCCGGTTGGAG 627
|||||
Db 19004 TCCCCAGCCGCGAAGCGAGAGGGCGATGTTTCAGGCTGCAGGGCANGCCGGTGGAG 19063
QY 628 CAGCAGCCACCACCGAGAGACCCGAGACAGTGTGGCGCGATGCAGTCTGAGCCTGGAGC 687
|||||
Db 19064 CAGCAGCCACCACCGAGAGACCCGAGACAGTGTGGCGCGATGCAGTCTGAGCCTGGAGC 19123
QY 688 GAGGAGCCATGTCCTCCCGCAGAGATGGCTCAGACCCCGTGATTCGATGCCCTCGA 747
|||||
Db 19124 GAGGAGCCANN 19183
QY 748 GACCAGGCTCCCGCAAGCAGCAGCAGCAGCAGCAGCGCTGCTTC 798
Db 19184 NNGC 19234

RESULT 4
AC125993 173007 bp DNA linear HTG 23-JUL-2002
LOCUS Rattus norvegicus clone CH230-74L11, *** SEQUENCING IN PROGRESS
DEFINITION *** 59 unordered pieces.
AC125993
AC125993.1 GI:21671835
VERSION HTG, HTGS_PHASE1.
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS 1 (bases 1 to 173007)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barberia,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,

TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL
TITLE
AUTHORS
REFERENCE
JOURNAL
TITLE
AUTHORS
REFERENCE
JOURNAL
COMMENT

Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 173007)
Worley,K.C.
Direct Submission
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 173007)
Worley,K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GC12
Center clone name: CH230-74L11
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 119750 bases at least Q40
Consensus quality: 129093 bases at least Q30
Consensus quality: 135173 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 59 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1528 1527: contig of 1527 bp in length
1628 1627: gap of unknown length
1628 2780: contig of 1153 bp in length
2781 2880: gap of unknown length
2881 3908: contig of 1028 bp in length
3909 4008: gap of unknown length
4009 5112: contig of 1104 bp in length
5113 5212: gap of unknown length
5213 6676: contig of 1464 bp in length
6677 6776: gap of unknown length
6777 8437: contig of 1661 bp in length
8438 8537: gap of unknown length
8538 9579: contig of 1042 bp in length
9580 9679: gap of unknown length
9680 11551: contig of 1872 bp in length
11552 11651: gap of unknown length
11652 13095: contig of 1444 bp in length
13096 13195: gap of unknown length
13196 14531: contig of 1336 bp in length
14532 14631: gap of unknown length
14632 15772: contig of 1141 bp in length
15773 15872: gap of unknown length
15873 17701: contig of 1829 bp in length
17702 17801: gap of unknown length
17802 18902: contig of 1101 bp in length
18903 19002: gap of unknown length
19003 21098: contig of 2096 bp in length
21099 22579: gap of unknown length
22580 22679: contig of 1381 bp in length
22680 23984: gap of unknown length
23984: contig of 1305 bp in length

KEYWORDS

HTG; HTGS_PHASE1.

SOURCE

Norway rat.

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE

1 (bases 1 to 89551)

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Hollaway,C., Hollins,B.,
Homs,I.F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE

Direct Submission

JOURNAL

2 (bases 1 to 89551)

AUTHORS

Worley,K.C.

TITLE

Direct Submission

JOURNAL

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 89551)

AUTHORS

Worley,K.C.

TITLE

Direct Submission

JOURNAL

Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KAMV

Center clone name: CH230-69F8

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 42669 bases at least Q40
Consensus quality: 46367 bases at least Q30
Consensus quality: 48722 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1151: contig of 1151 bp in length
1152 1251: gap of unknown length
1252 2322: contig of 1071 bp in length
2323 2422: gap of unknown length
2423 3692: contig of 1270 bp in length
3693 3792: gap of unknown length
3793 5136: contig of 1344 bp in length
5137 5236: gap of unknown length
5237 6325: contig of 1089 bp in length
6326 6425: gap of unknown length
6426 8049: contig of 1624 bp in length
8050 8149: gap of unknown length
8150 9419: contig of 1270 bp in length
9420 9519: gap of unknown length
9520 10641: contig of 1122 bp in length
10642 10741: gap of unknown length
10742 11933: contig of 1192 bp in length
11934 12033: gap of unknown length
12034 13362: contig of 1329 bp in length
13363 13462: gap of unknown length
13463 14779: contig of 1317 bp in length
14780 14879: gap of unknown length
14880 16451: contig of 1572 bp in length
16452 16551: gap of unknown length
16552 17915: contig of 1364 bp in length
17916 18015: gap of unknown length
18016 19443: contig of 1428 bp in length
19444 19543: gap of unknown length
19544 20626: contig of 1083 bp in length
20627 20726: gap of unknown length
20727 22100: contig of 1374 bp in length
22101 22200: gap of unknown length
22201 23352: contig of 1052 bp in length
23353 23352: gap of unknown length
23354 24945: contig of 1593 bp in length
24946 25045: gap of unknown length
25046 26886: contig of 1841 bp in length
26887 26986: gap of unknown length
26987 28215: contig of 1229 bp in length
28216 28315: gap of unknown length
28316 29904: contig of 1589 bp in length
29905 30004: gap of unknown length
30005 31803: contig of 1799 bp in length
31804 31903: gap of unknown length
31904 33938: contig of 2035 bp in length
33939 34038: gap of unknown length
34039 35078: gap of unknown length
35079 35178: contig of 1040 bp in length
35179 37352: gap of unknown length
37353 37452: gap of unknown length
37453 39726: contig of 2274 bp in length
39727 39826: gap of unknown length
39827 41555: contig of 1729 bp in length
41556 41655: gap of unknown length
41656 43271: contig of 1616 bp in length
43272 43371: gap of unknown length
43372 44785: contig of 1414 bp in length
44786 44885: gap of unknown length
44886 46567: contig of 1682 bp in length

*	46568	46667:	gap of unknown length
*	46668	49449:	contlg of 2782 bp in length
*	49450	49549:	gap of unknown length
*	49550	51783:	contlg of 2234 bp in length
*	51784	51883:	gap of unknown length
*	51884	54046:	contlg of 2163 bp in length
*	54047	54146:	gap of unknown length
*	54147	56182:	contlg of 2036 bp in length
*	56183	56282:	gap of unknown length
*	56283	57934:	contlg of 1652 bp in length
*	57935	58034:	gap of unknown length
*	58035	59825:	contlg of 1791 bp in length
*	59826	59925:	gap of unknown length
*	59926	61535:	contlg of 1610 bp in length
*	61536	61635:	gap of unknown length
*	61636	63473:	contlg of 1838 bp in length
*	63474	63573:	gap of unknown length
*	63574	66376:	contlg of 2803 bp in length
*	66377	66476:	gap of unknown length
*	66477	68641:	contlg of 2165 bp in length
*	68642	68741:	gap of unknown length
*	68742	70180:	contlg of 1439 bp in length
*	70181	70280:	gap of unknown length
*	70281	72432:	contlg of 2152 bp in length
*	72433	72532:	gap of unknown length
*	72533	73917:	contlg of 1385 bp in length
*	73918	74017:	gap of unknown length
*	74018	77137:	contlg of 3120 bp in length
*	77138	77237:	gap of unknown length
*	77238	79997:	contlg of 2760 bp in length
*	79998	80097:	gap of unknown length
*	80098	82859:	contlg of 2762 bp in length
*	82860	82959:	gap of unknown length
*	82960	85680:	contlg of 2721 bp in length
*	85681	85780:	gap of unknown length
*	85781	89551:	contlg of 3771 bp in length

BASE COUNT	ORIGIN
23140 a	19260 c
19626 g	22768 t
4757 others	

Query Match	45.88;	Score 585.2;	DB 2;	Length 89551;
Best Local Similarity	87.68;	Pred. No. 2.1e-110;		
Matches 664;	Conservative 0;	Mismatches 88;	Indels 6;	Gaps 2

[illegible]

Db	39327	TCCTGGCGCACCCGTGCCCCCTGTACTGCCCCGTGACCTTCGGTGGCCTCTCCTCGCTG	39268
QY	466	GAGTTGCGGAGCGCAGGCAGACACCCACGAAGGAGAGGGAGCCCCGCATCTCGGG	525
Db	39267	GAGTTGCGGGGAGCAGGCAGACGCCACAGGAAGGAGAGGGAGACCCGACACCGGG	39208
QY	526	ACCCGGGAACCGGAGCCGAGAGAGAGTGGCCGCGAGAAAGCGTCCCCACGCCGGAAGC	585
Db	39207	ACCCGGGAACCCGAGCCCGGAGAGAGTGGCAGTGATGAAGCAGTCCCCACGCCGAGAGC	39148
QY	586	GAGGAGGGCGATGTTCAAGCTGCAGGGCGACGCCGGGTGGGAGCAGCCACCACCGGAG	645
Db	39147	GAGGAGGGCGACGTCACAGCTGAAGGGCAGGATG--GGCAGGAGCAGCCACC GGCGAG	39091
QY	646	GACCGGAACAGTGTGGCGCGGATGCAGTCTGAGCCTGGGAGCGGAGGCCATGTCTGCC	705
Db	39090	GACCCGACAGTGTGGCGCGGATGCAGTCTGAGCCCCGGAGTGAAGGACCACTCTCTGT	39031
QY	706	GCAGAGATGGCTCAGGACCCCCGGTGATTCCGATGCCCTCGAGACCAGGCCCTCCCGCA	765
Db	39030	GTCGAGATGGCTCAGGACCCCAAGTACGTGGCTGCCCTTAGAGACCGGGCCTCCCAAG	38971
QY	766	AGCAGGAGCAGGACAAAGAGCGCCTGCGTTTCCAGTT	803
Db	38970	AGCACTGAGCAGGACAAAGAGCGCCTGCGTTTCCAGGT	38933

RESULT	6
AC126519/c	
LOCUS	
DEFINITION	
AC126519	169908 bp DNA linear HTG 24-JUL-2002
Rattus norvegicus clone CH230-159N5,	*** SEQUENCING IN PROGRESS
***,	60 unordered pieces.

ACCESSION	AC126519
VERSION	AC126519.1
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Rattus norvegicus
ORGANISM	Rattus norvegicus

REFERENCE AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 169908)

Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burreli,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagay,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newtonson,J., Newtonson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwouu,G., Otagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,

Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatok, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Unpublished
Direct Submission
2 (bases 1 to 169908)
Worley, K.C.

Direct Submission
Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 169908)
Worley, K.C.

Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GZCZ
Center clone name: CH230-159N5

----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 125808 bases at least Q40
Consensus quality: 132708 bases at least Q30
Consensus quality: 136844 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1575 1574: contig of 1574 bp in length
1675 1674: gap of unknown length
2675 2679: contig of 1005 bp in length
2680 2779: gap of unknown length
2780 3878: contig of 1099 bp in length
3879 3978: gap of unknown length
3979 5423: contig of 1445 bp in length
5424 5523: gap of unknown length
5524 6788: contig of 1265 bp in length
6789 6888: gap of unknown length
6889 8285: contig of 1397 bp in length
8286 8385: gap of unknown length
8386 9590: contig of 1205 bp in length
9591 9690: gap of unknown length
9691 10999: contig of 1309 bp in length
11000 11099: gap of unknown length
11100 12498: contig of 1399 bp in length
12499 12598: gap of unknown length
12599 14147: contig of 1549 bp in length
14148 14247: gap of unknown length
14248 16266: contig of 2019 bp in length
16267 16366: gap of unknown length
16367 18233: contig of 1867 bp in length
18234 18333: gap of unknown length

18334 19761: contig of 1428 bp in length
19762 19861: gap of unknown length
19862 21637: contig of 1776 bp in length
21638 21737: gap of unknown length
21738 23143: contig of 1406 bp in length
23144 23243: gap of unknown length
23244 25296: contig of 2053 bp in length
25297 25396: gap of unknown length
25397 26533: contig of 1137 bp in length
26534 26633: gap of unknown length
26634 28278: contig of 1645 bp in length
28279 28378: gap of unknown length
28379 30343: contig of 1965 bp in length
30344 30443: gap of unknown length
30444 32121: contig of 1678 bp in length
32122 32221: gap of unknown length
32222 34582: contig of 2361 bp in length
34583 34682: gap of unknown length
34683 35980: contig of 1298 bp in length
35981 36080: gap of unknown length
36081 37394: contig of 1314 bp in length
37395 37494: gap of unknown length
37495 38968: contig of 1474 bp in length
38969 39068: gap of unknown length
39069 40199: contig of 1131 bp in length
40200 40299: gap of unknown length
40300 41804: contig of 1505 bp in length
41805 41904: gap of unknown length
41905 43569: contig of 1665 bp in length
43570 43669: gap of unknown length
43670 45680: contig of 2011 bp in length
45681 45780: gap of unknown length
45781 48215: contig of 2435 bp in length
48216 48315: gap of unknown length
48316 51068: contig of 2753 bp in length
51069 51168: gap of unknown length
51169 53666: contig of 2498 bp in length
53667 53766: gap of unknown length
53767 55832: contig of 2066 bp in length
55833 55932: gap of unknown length
55933 57560: contig of 1628 bp in length
57561 57660: gap of unknown length
57661 59904: contig of 2244 bp in length
59905 60004: gap of unknown length
60005 62974: contig of 2970 bp in length
62975 63074: gap of unknown length
63075 66491: contig of 3417 bp in length
66492 66591: gap of unknown length
66592 68271: contig of 1680 bp in length
68272 68371: gap of unknown length
68372 70762: contig of 2391 bp in length
70763 70862: gap of unknown length
70863 73263: contig of 2401 bp in length
73264 73363: gap of unknown length
73364 76903: contig of 3540 bp in length
76904 77003: gap of unknown length
77004 79444: contig of 2441 bp in length
79445 79544: gap of unknown length
79545 81957: contig of 2413 bp in length
81958 82057: gap of unknown length
82059 85214: contig of 3157 bp in length
85215 85314: gap of unknown length
85315 87778: contig of 2464 bp in length
87779 87878: gap of unknown length
87879 90897: contig of 3019 bp in length
90898 90997: gap of unknown length
90999 96387: contig of 5390 bp in length
96388 96487: gap of unknown length
96488 100595: contig of 4108 bp in length
100596 100695: gap of unknown length
100696 104991: contig of 4296 bp in length
104992 105091: gap of unknown length
105092 109386: contig of 4295 bp in length

*	109387	109486:	gap of unknown length
*	109487	113734:	contig of 4248 bp in length
*	113735	118834:	gap of unknown length
*	113835	116884:	contig of 3050 bp in length
*	116885	116984:	gap of unknown length
*	116985	121008:	contig of 4024 bp in length
*	121009	121108:	gap of unknown length
*	121109	124634:	contig of 3526 bp in length
*	124635	124734:	gap of unknown length
*	124735	130307:	contig of 5573 bp in length
*	130308	130407:	gap of unknown length

Query Match	35.48;	Score 452.4;	DB 2;	Length 169908;
Best Local Similarity	87.58;	Pred. No. 4.9e-83;		
Matches 530; Conservative	0;	Mismatches 71;	Indels 5;	Gaps 3;

QY	199	CAGCTCATGGCCCTGCTGTGCGGATGGGTCCCGGTCG-GTCAGCAGCCGTGACGCTGC	257
Db	10971	CACTTCATGGCCCTTGCTGTGCGAATGGGTCCCGGTCGCGTCAGCAGCCGCGACGCTGC	10912
QY	258	GCTGCAGGTGAACCCGCGCGCGAGCGCTTCGTGTCAGTGTTCACCTCGGGCGCCGACGCT	317
Db	10911	GGTGCAGGTGAACCCGCGCGCGAGCTTCGTGTCAGTGTTCGCTCGGGCGCCGACACT	10852
QY	318	GCAGCTTCAGGGGTGCGGAGCCAGCCCGGACGCCCGATCGGGTTCCTGTCAACCCCGTGG	377
Db	10851	GCAGCTTCGACCGGCGCGCGAGCCAGCCCGGACGCCCGGCTGTTCTCTGCCAACCCCGCAG	10792
QY	378	CCACGCCCGCGCGCGGAGATCCCCGCGATCCTGGCAGACCCGTAGCCCCGTTCTCTCCCT	437
Db	10791	CCCCGCCAGGGCGCGGAGACCCCCGCGATCCTGGCGCACCGTCGCCCTGTACTCGCCCT	10732
QY	438	GACCTTCTGTGGCTCTCTCCTCCTCACTGAGAGTTGCGGGAGGACGACAGCACCCACGAA	497
Db	10731	GACCTTCGGTGGCTCTCTCCTCCTCGCTGAGAGTTGCGGGGACAGGACAGCCCCACGAA	10672
QY	498	GGGAGAGGGGAGCCCGGCATCCTCGGGGACCCGGGAACCGGAGCCGAGAGGTTGGCCGC	557
Db	10671	-GGAGAGGGGAGACCGGCACCCACGGGGACCCGGGAACCCGAGCCGGGAGAGGTTGGCAGT	10613
QY	558	GAGGAAGCGGTTCCCCCAGCCGCGGAAGCGAGGAGGGCGATGTTCAAGCTGCAGGGCAGGC	617
Db	10612	GATGAAGCAGTCCCCCAGCCGCGCAGAGCGAGGAGGGCGACGTCCAGGCTGAAGGGCAGGA	10553
QY	618	CGGGTGGGAGCAGCAGCCACACCGGAGGAGCCGGAACAGTGTGGCGGCGATGACAGTCTGA	677
Db	10552	TG--GGCAGGAGCAGCACCGCGGGAGGAGCCCGACAGTGTGGCGGCGATGACAGTCTGA	10496
QY	678	GCCTGGGAGCGAGGAGCCATGCTGCGCGCAGAGATGGCTCAGGACCCCGGTGATTCCGA	737
Db	10495	GCCCGGGAGTGAGGAGCCACCTCTGCTGTGAGATGGCTCAGGACCCCACTGACGTGGC	10436
QY	738	TGCCCCTCGAGACCAAGGCTCCCGCAAAAGCACGGAGCAGGACAGCAAGAGCGGCTGCGTTT	797
Db	10435	TGCTCTAGAGACCGGGGCTTCCCAAGAGCACTGAGCAGGACAAAGAGCGGCTGCGTTT	10376
QY	798	CCAGTT 803	
Db	10375	CCAGTT 10370	

RESULT 7	AC108848	LOCUS	DEFINITION	AC108848	212848 bp	DNA	linear	HTG 28-APR-2002
AC108848								
AC108848								
AC108848								
AC108848.2								
GI:20336129								
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.								
Mus musculus.								
Mus musculus								
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 212848)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-300K5
Unpublished

Birfen, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bokhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Huime, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
Landers, T., Lehoczeky, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Piere, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosettl, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE	REFERENCE	JOURNAL	AUTHORS
1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20
21	22	23	24
25	26	27	28
29	30	31	32
33	34	35	36
37	38	39	40
41	42	43	44
45	46	47	48
49	50	51	52
53	54	55	56
57	58	59	60
61	62	63	64
65	66	67	68
69	70	71	72
73	74	75	76
77	78	79	80
81	82	83	84
85	86	87	88
89	90	91	92
93	94	95	96
97	98	99	100

Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 212848)

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R.,
 Landers, T., Lechoczky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 28, 2002 this sequence version replaced q1:18450109.

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)

----- Genome Center
http://flp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information -----

Center project name: L21014

Center clone name: 300_K_5

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 1008 of reads

Chemistry: Dye-terminator Blg Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 200069 bases at least Q40

Consensus quality: 206302 bases at least Q30
Consensus quality: 208218 bases at least Q20
Insert size: 209348; sum-of-contigs
Quality coverage: 5.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 36 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

1 641: contig of 641 bp in length
642 741: gap of 100 bp
742 1427: contig of 686 bp in length
1428 1527: gap of 100 bp
1528 3170: contig of 1643 bp in length
3171 3270: gap of 100 bp
3271 4396: contig of 1126 bp in length
4397 4496: gap of 100 bp
4497 5909: contig of 1413 bp in length
5910 6009: gap of 100 bp
6010 7583: contig of 1574 bp in length
7584 7683: gap of 100 bp
7684 8975: contig of 1292 bp in length
8976 9075: gap of 100 bp
9076 10257: contig of 1182 bp in length
10258 10357: gap of 100 bp
10358 11608: contig of 1251 bp in length
11609 11708: gap of 100 bp
11709 13289: contig of 1581 bp in length
13290 13389: gap of 100 bp
13390 15335: contig of 1946 bp in length
15336 15435: gap of 100 bp
15436 16250: contig of 815 bp in length
16251 16350: gap of 100 bp
16351 17669: contig of 1319 bp in length
17670 17769: gap of 100 bp
17770 19907: contig of 2138 bp in length
19908 20007: gap of 100 bp
20008 21470: contig of 1463 bp in length
21471 21570: gap of 100 bp
21571 23086: contig of 1516 bp in length
23087 23186: gap of 100 bp
23187 24723: contig of 1537 bp in length
24724 24823: gap of 100 bp
24824 27746: contig of 2923 bp in length
27747 27846: gap of 100 bp
27847 29873: contig of 2027 bp in length
29874 29973: gap of 100 bp
29974 33212: contig of 3239 bp in length
33213 33312: gap of 100 bp
33313 61732: contig of 28420 bp in length
61733 61832: gap of 100 bp
61833 63792: contig of 1960 bp in length
63793 63892: gap of 100 bp
63893 67231: contig of 3339 bp in length
67232 67331: gap of 100 bp
67332 70922: contig of 3591 bp in length
70923 71022: gap of 100 bp
71023 76930: contig of 5908 bp in length
76931 77030: gap of 100 bp
77031 82623: contig of 5593 bp in length
82624 82723: gap of 100 bp
82724 87875: contig of 5152 bp in length
87876 87975: gap of 100 bp
87976 93774: contig of 5799 bp in length
93775 93874: gap of 100 bp
93875 104205: contig of 10331 bp in length
104206 104305: gap of 100 bp
104306 119104: contig of 14799 bp in length

* 119105 119204: gap of 100 bp
* 119205 132299: contig of 13095 bp in length
* 132300 132399: gap of 100 bp
* 132400 148970: contig of 16571 bp in length
* 148971 149070: gap of 100 bp
* 149071 165714: contig of 16644 bp in length
* 165715 165814: gap of 100 bp
* 165815 183369: contig of 17555 bp in length
* 183370 183469: gap of 100 bp
* 183470 210291: contig of 26822 bp in length
* 210292 210391: gap of 100 bp
* 210392 212848: contig of 2457 bp in length.

FEATURES

source

1. 212848
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RP23-300K5"
/clone_lib="RP23-300K5"
1. 641
/note="assembly_fragment"
742. 1427
/note="assembly_fragment"
1528. 3170
/note="assembly_fragment"
3271. 4396
/note="assembly_fragment"
4497. 5909
/note="assembly_fragment"
6010. 7583
/note="assembly_fragment"
7684. 8975
/note="assembly_fragment"
9076. 10257
/note="assembly_fragment"
10358. 11608
/note="assembly_fragment"
11709. 13289
/note="assembly_fragment"
13390. 15335
/note="assembly_fragment"
15436. 16250
/note="assembly_fragment"
16351. 17669
/note="assembly_fragment"

Query Match 28.2%; Score 360; DB 2; Length 212848;
Best Local Similarity 91.7%; Pred. No. 5.4e-64;
Matches 429; Conservative 0; Mismatches 21; Indels 18; Gaps 4;

QY 336 AGCCAGCCCCGACGCCGATCGGGTCTCTCAACCCCGTGCCACGCCGCGCGGAG 395
DB 93875 AGCCAGCCCCGACGCCGATCGGGTCTCTCAACCCCGTGCCACGCCGCGGAG 93934
QY 396 ATCCCGCGATCCTGGCAGACCGTAGCCCGTCTCTGCTGACCTTGTGCGCTCTC 455
DB 93935 ATCCCGCGATCCTGGCAGACCGTAGCCCGTCTCTGCTGACCTTGTGCGCTCTC 93990
QY 456 CTCCTCACTGAGGTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 515
DB 93991 CTCCTCACTGAGGTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 94049
QY 516 ATCCTCGGGGACCCCGGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 575
DB 94050 ATCCTCGGGGACCCCGGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 94109
QY 576 GCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 635
DB 94110 GCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 94169
QY 636 ACCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 695
DB 94170 ACCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 94229

QY 696 ATGTCCTGCCGAGAGATGGCTCAGAGACCCGGTGATTCGGATCCCTCGAGACCAGCC 755
 Db 94230 GTGTCCTGCTGAGAGATGGCTCAGAGACCCGGTGACTCGGATGC-----CC 94276
 QY 756 CTCCCGCAAGCAGCAGGAGAGACAGCCGCTGCTTCAGTT 803
 Db 94277 CTCCCGCAAGCAGCAGGAGAGACAGCAGCAGCTGCTGCTTCAGGT 94324

RESULT 8
 AC108848/c 212848 bp DNA linear HTG 28-APR-2002
 LOCUS Mus musculus clone RP23-300K5, WORKING DRAFT SEQUENCE, 36 ordered
 DEFINITION pieces.
 AC108848
 AC108848.2 GI:20336129
 VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.

REFERENCE 1
 AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouknight, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
 Landers, T., Lehoczký, J., Levine, R., Liu, G., Maclean, C.,
 MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Plerre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Title Mus musculus, clone RP23-300K5
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 212848)

TITLE Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome
 JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 212848)
 AUTHORS

Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Bouknight, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
 Landers, T., Lehoczký, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
 Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Plerre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 COMMENT

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 28, 2002 this sequence version replaced g1:18450109.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L21014

Center clone name: 300_K_5

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 200069 bases at least Q40

Consensus quality: 206302 bases at least Q30

Consensus quality: 208218 bases at least Q20

Insert size: 209348; sum-of-contigs

Quality coverage: 5.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 36 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

of the gaps between them are based on estimates that have

provided by the submitter.

This sequence will be replaced

by the finished sequence as soon as it is available and

the accession number will be preserved.

1 642 741: gap of 100 bp

742 1427: contig of 686 bp in length

1428 1527: gap of 100 bp

1528 3170: contig of 1643 bp in length

3171 3270: gap of 100 bp

3271 4396: contig of 1126 bp in length

4397 4496: gap of 100 bp

4497 5909: contig of 1413 bp in length

5910 6009: gap of 100 bp

6010 7583: contig of 1574 bp in length

7584 7683: gap of 100 bp

7684 8975: contig of 1292 bp in length

8976 9075: gap of 100 bp

9076 10257: contig of 1182 bp in length

10258 10357: gap of 100 bp

10358 11608: contig of 1251 bp in length

11609 11708: gap of 100 bp

11709 13289: contig of 1581 bp in length

13290 13389: gap of 100 bp

13390 15335: contig of 1946 bp in length

15336 15435: gap of 100 bp

15436 16250: contig of 815 bp in length

16251 16350: gap of 100 bp

16351 17669: contig of 1319 bp in length

17670 17769: gap of 100 bp

17770 19907: contig of 2138 bp in length

19908 20007: gap of 100 bp

20008 21470: contig of 1463 bp in length

21471 21570: gap of 100 bp

21571 23086: contig of 1516 bp in length

23087 23186: gap of 100 bp

23187 24723: contig of 1537 bp in length

24724 24823: gap of 100 bp

24824 27746: contig of 2923 bp in length

27747 27846: gap of 100 bp

27847 29873: contig of 2027 bp in length

29874 29973: gap of 100 bp

* 29974 33212: contig of 3239 bp in length
* 33213 33312: gap of 100 bp
* 33313 61732: contig of 28420 bp in length
* 61733 61832: gap of 100 bp
* 61833 63792: contig of 1960 bp in length
* 63793 63892: gap of 100 bp
* 63893 67231: contig of 3339 bp in length
* 67232 67331: gap of 100 bp
* 67332 70922: contig of 3591 bp in length
* 70923 71022: gap of 100 bp
* 71023 76930: contig of 5908 bp in length
* 76931 77030: gap of 100 bp
* 77031 82623: contig of 5593 bp in length
* 82624 82723: gap of 100 bp
* 82724 87875: contig of 5152 bp in length
* 87876 87975: gap of 100 bp
* 87976 93774: contig of 5799 bp in length
* 93775 93874: gap of 100 bp
* 93875 104205: contig of 10331 bp in length
* 104206 104305: gap of 100 bp
* 104306 119104: contig of 14799 bp in length
* 119105 119204: gap of 100 bp
* 119205 132299: contig of 13095 bp in length
* 132300 132399: gap of 100 bp
* 132400 148970: contig of 16571 bp in length
* 148971 149070: gap of 100 bp
* 149071 165714: contig of 16644 bp in length
* 165715 165814: gap of 100 bp
* 165815 183369: contig of 17555 bp in length
* 183370 183469: gap of 100 bp
* 183470 210291: contig of 26822 bp in length
* 210292 210391: gap of 100 bp
* 210392 212848: contig of 2457 bp in length.

FEATURES

source
1. 212848
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-300K5"
/clone_lib="RPCI-23 Female Mouse BAC"
misc_feature
1. 641
/note="assembly_fragment"
742. 1427
/note="assembly_fragment"
misc_feature
1528. 3170
/note="assembly_fragment"
misc_feature
3271. 4396
/note="assembly_fragment"
misc_feature
4497. 5909
/note="assembly_fragment"
misc_feature
6010. 7583
/note="assembly_fragment"
misc_feature
7684. 8975
/note="assembly_fragment"
misc_feature
9076. 10257
/note="assembly_fragment"
misc_feature
10358. 11608
/note="assembly_fragment"
misc_feature
11709. 13289
/note="assembly_fragment"
misc_feature
13390. 15335
/note="assembly_fragment"
misc_feature
15436. 16250
/note="assembly_fragment"
misc_feature
16351. 17669
/note="assembly_fragment"

Query Match 23.5%; Score 299.8; DB 2; Length 212848;
Best Local Similarity 72.6%; Pred. No. 1.4e-51;
Matches 310; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 3 GCGGGGCGAGCGGCGGACCGCATGTTCCTCCGCGAGCAGCTTCACCCCTGCCCCGCA 62
Db 88270 GCGGGGCGAGCGGCGGACCGCATGTTCCTCCGCGAGCAGCTTCACCCCTGCCCCGCA 88211

QY 63 TCCTTATCCGCGAGCCACCAAGCCGGGATGGCTGAGAGTTCCGAGCCAGGGCTGCCG 122
Db 88210 TCCTTATCCGCGAGCCACCAAGCCGGGATGGCTGAGAGTTCCGAGCCAGGGCTGCCG 88151
QY 123 ACCCGCGCCCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGGCCGGAGTACGTCA 182
Db 88150 ACCCGCGCCCCCTCTCTCTCTCCCGGCTACAGACAGCTCATGGCCGGAGTACGTCA 88091
QY 183 CAGCCACGAGCGGGCAGAGCTCATGGCCCTGCTGTCCGGATGGTCCCGGCTGAC 242
Db 88090 CCGCCACGAGCGGGCAGAGCTCATGGCCCTGCTGTCCGGATGGTCCCGGCTGAC 88031
QY 243 CAGCCGTGACGCTGCGGTGAGTGAACCCGCGCGCCGCTCGGTGACGTTCAC 302
Db 88030 CAGCCGTGACGCTGCGGTGAGTGAACCCGCGCGCCGCTCGGTGACGTTCAC 87971
QY 303 CCGGCGCGCGCAGCGCTGACGCTGAGGCTGCCGAGCCAGCCCGCAGCTCGGTTTC 362
Db 87970 NNN 87911
QY 363 CTGTCAACCCCGTGCGCCACCGCGCGCGGAGATCCCGCGCATCTGCGAGACCGTAGC 422
Db 87910 NNN 87851
QY 423 CCCGTTTC 429
Db 87850 ACAGTTTC 87844

RESULT 9
AC127083
LOCUS
DEFINITION
Rattus norvegicus clone CH230-69F8, *** SEQUENCING IN PROGRESS ***
48 unordered pieces.
AC127083
AC127083.1 GI:21743720
VERSION
HTG; HTGS_PHASE1.
KEYWORDS
SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
1 (bases 1 to 89551)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Delaey,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Joudah,S.,
Jacobson,B., Jia,Y., Johnson,R., Jolyvet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters, L., Pickens, R., Primus, E., Pu, L.L., Qules, M., Ren, Y.,
Rives, M., Rojas, A., Rojubenokan, I., Rolfe, M., Ruiz, S., Saverly, G.,
Soderer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonatke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

REFERENCE
JOURNAL
TITLE
AUTHORS
JOURNAL
2 (bases 1 to 89551)
Unpublished
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 89551)
Worley, K.C.

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: KAMV
Center clone name: CH230-69F8
Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 42669 bases at least Q40
Consensus quality: 46367 bases at least Q30
Consensus quality: 48722 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1152 1151: contig of 1151 bp in length
1252 1251: gap of unknown length
2322 2322: contig of 1071 bp in length
2323 2422: gap of unknown length
2423 3692: contig of 1270 bp in length
3693 3793: gap of unknown length
3793 5137: contig of 1344 bp in length
5137 5237: gap of unknown length
5237 6326: contig of 1089 bp in length
6326 6426: gap of unknown length
6426 8050: contig of 1624 bp in length
8050 8150: gap of unknown length
8150 9420: contig of 1270 bp in length
9420 9520: gap of unknown length
9520 10642: contig of 1122 bp in length
10642 10742: gap of unknown length
10742 11934: contig of 1192 bp in length
11934 12034: gap of unknown length
12034 13362: contig of 1329 bp in length
13362 13463: gap of unknown length
13463 14780: contig of 1317 bp in length
14780 14879: gap of unknown length

14880 16451: contig of 1572 bp in length
16452 16551: gap of unknown length
16552 17915: contig of 1364 bp in length
17916 18015: gap of unknown length
18016 19443: contig of 1428 bp in length
19444 19543: gap of unknown length
19544 20626: contig of 1083 bp in length
20627 20726: gap of unknown length
20727 22100: contig of 1374 bp in length
22101 22200: gap of unknown length
22201 23252: contig of 1052 bp in length
23253 23352: gap of unknown length
23353 24945: contig of 1593 bp in length
24946 25045: gap of unknown length
25046 26886: contig of 1841 bp in length
26887 26986: gap of unknown length
26987 28215: contig of 1229 bp in length
28216 28315: gap of unknown length
28316 29904: contig of 1589 bp in length
29905 30004: gap of unknown length
30005 31803: contig of 1799 bp in length
31804 31903: gap of unknown length
31904 33938: contig of 2035 bp in length
33939 34038: gap of unknown length
34039 35078: contig of 1040 bp in length
35079 35178: gap of unknown length
35179 37352: contig of 2174 bp in length
37353 37452: gap of unknown length
37453 39726: contig of 2274 bp in length
39727 39826: gap of unknown length
39827 41555: contig of 1729 bp in length
41556 41655: gap of unknown length
41656 43271: contig of 1616 bp in length
43272 43371: gap of unknown length
43372 44785: contig of 1414 bp in length
44786 44885: gap of unknown length
44886 46567: contig of 1682 bp in length
46568 46667: gap of unknown length
46668 49449: contig of 2782 bp in length
49450 49549: gap of unknown length
49550 51783: contig of 2234 bp in length
51784 51883: gap of unknown length
51884 54046: contig of 2163 bp in length
54047 54146: gap of unknown length
54147 56182: contig of 2036 bp in length
56183 56282: gap of unknown length
56283 57934: contig of 1652 bp in length
57935 58034: gap of unknown length
58035 59825: contig of 1791 bp in length
59826 59925: gap of unknown length
59926 61535: contig of 1610 bp in length
61536 61635: gap of unknown length
61636 63473: contig of 1838 bp in length
63474 63573: gap of unknown length
63574 66376: contig of 2803 bp in length
66377 66476: gap of unknown length
66477 68641: contig of 2165 bp in length
68642 70180: contig of 1439 bp in length
70181 70280: gap of unknown length
70281 72432: contig of 2152 bp in length
72433 72532: gap of unknown length
72533 73917: contig of 1385 bp in length
73918 74017: gap of unknown length
74018 77137: contig of 3120 bp in length
77138 77237: gap of unknown length
77238 79997: contig of 2760 bp in length
79998 80097: gap of unknown length
80098 82859: contig of 2762 bp in length
82859 82959: gap of unknown length
82960 85680: contig of 2721 bp in length
85681 85780: gap of unknown length
85781 89551: contig of 3771 bp in length.


```
repeat_region /rpt_family="L1"
10850. .11026
repeat_region /rpt_family="L2"
11461. .17576
repeat_region /rpt_family="L1"
17680. .17758
repeat_region /rpt_family="Mariner"
18412. .18629
repeat_region /rpt_family="MIR"
19942. .20121
repeat_region /rpt_family="MERL_type"
20456. .20629
repeat_region /rpt_family="(TA)n"
20645. .20757
repeat_region /rpt_family="L1"
20758. .21114
repeat_region /rpt_family="MALR"
22734. .23101
repeat_region /rpt_family="MALR"
23102. .24334
repeat_region /rpt_family="L1"
24340. .24720
repeat_region /rpt_family="L1"
24721. .25077
repeat_region /rpt_family="L1"
25078. .25115
repeat_region /rpt_family="(TA)n"
25116. .25599
repeat_region /rpt_family="L1"
25600. .25625
repeat_region /rpt_family="AT_rich"
26058. .26331
repeat_region /rpt_family="L2"
26345. .26475
repeat_region /rpt_family="L2"
27007. .27085
repeat_region /rpt_family="GA-rich"
27113. .27146
repeat_region /rpt_family="(CA)n"
27479. .27605
repeat_region /rpt_family="MERL_type?"
27924. .28079
repeat_region /rpt_family="MERL_type"
29818. .30292
repeat_region /rpt_family="MERL_type"
31387. .31688
repeat_region /rpt_family="Alu"
32916. .32981
repeat_region /rpt_family="CT-rich"
32982. .33240
repeat_region /rpt_family="Alu"
33241. .33252
repeat_region /rpt_family="CT-rich"
34110. .34142
repeat_region /rpt_family="AT_rich"
35992. .36014
repeat_region /rpt_family="(TAAA)n"
36034. .36187
repeat_region /rpt_family="Retroviral"
36188. .36481
repeat_region /rpt_family="Alu"
36482. .36514
repeat_region /rpt_family="Retroviral"
36617. .36835
repeat_region /rpt_family="L1"
36838. .36905
repeat_region /rpt_family="(TC)n"
36906. .36939
repeat_region /rpt_family="(TA)n"
36945. .37725
repeat_region /rpt_family="L1"
37736. .38050
repeat_region /rpt_family="Alu"
```

```
repeat_region 38070. .38557
/rpt_family="MER4-group"
repeat_region 38558. .38976
/rpt_family="L2"
repeat_region 39061. .39116
/rpt_family="(TAAA)n"
repeat_region 39477. .39692
/rpt_family="MIR"
repeat_region 40301. .40339
/rpt_family="(TC)n"
repeat_region 40405. .40690
/rpt_family="Alu"
repeat_region 40713. .40816
/rpt_family="MERL_type"
misc_feature 41875. .42095
/note="match to EST AL046241 (NID:g5434325)"
```

Query Match 11.6%; Score 147.6; DB 9; Length 138890;
Best Local Similarity 73.7%; Pred. No. 3.2e-20;
Matches 202; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

```
QY 831 CTGCAAGACTGCAGAAATCCGGTGGAGAGAGCGCCTATGTGTGTGTGCAGGCAACCAG 890
|||||
Db 28657 CTGCAAGACTGTGACATCAATGGAGGTCCTATGTGTGTGTATAGACCTAA 28598

QY 891 TAAAGTGTACTTCAACAGCTTCTGCCAGTGTGTGAGAAATCCTACAAACCTTACAGAG 950
|||||
Db 28597 TAAGAACTACTTCAA--AGTTTACAGAACTGTGCAAGAGTCTTATTAACCTGTAACCAAG 28541

QY 951 TGGAGACATCACCCTGTCAAGTGTGTAAGAAGACTAGATGTGCTGCCAGTCAGATTTC 1010
|||||
Db 28540 GGGAGATATACAGCAATCAAGTGTGTAATAAAGTATGATGTGCTGCCAGTCAAGATTTC 28481

QY 1011 GCCACGTGGACCCCTAAAGCCCCCATCGGCAAGACTGTGTGGAGATGCAAGCAAAAC 1070
|||||
Db 28480 TCTACATAGACCCCTAAAGGCCCTATTGTCAAGATTGTGTGGAGATGCAAGCAAAAC 28421

QY 1071 GCCTGTCTCGGACAGCACTTCAGCTTCAAAATA 1104
|||||
Db 28420 ACCTATCTGTGTGAGCACTTTCAGCTTCAAAATA 28387
```

RESULT 11
AC037443/C 151961 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 2 clone RP11-77F17 map 2, WORKING DRAFT
DEFINITION
SEQUENCE, 10 unordered pieces.
ACCESSION AC037443
VERSION AC037443.2 GI:7960458
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 151961)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
TITILE Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
JOURNAL Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
REFERENCE Campopiano,A., Castle,A., Choepe,Y., Colangelo,M., Collins,S.,
AUTHORS Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

TITLE
JOURNAL
REFERENCE
AUTHORS

O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisanil, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 21, 2000 this sequence version replaced gi:7528172.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L9226
Center clone name: 77_F_17
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 147226 bases at least Q40
Consensus quality: 149421 bases at least Q30
Consensus quality: 150378 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 151061; sum-of-ctnigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 5.8 in Q20 bases; sum-of-ctnigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1451: contig of 1451 bp in length
* 1452 1551: gap of 100 bp
* 1552 3695: contig of 2144 bp in length
* 3696 3795: gap of 100 bp
* 3796 6752: contig of 2957 bp in length
* 6753 6852: gap of 100 bp

* 6853 18026: contig of 11174 bp in length
* 18027 18126: gap of 100 bp
* 18127 30673: contig of 12547 bp in length
* 30674 30773: gap of 100 bp
* 30774 44941: contig of 14168 bp in length
* 44942 45041: gap of 100 bp
* 45042 58179: contig of 13138 bp in length
* 58180 58279: gap of 100 bp
* 58280 74554: contig of 16275 bp in length
* 74555 74654: gap of 100 bp
* 74655 105703: contig of 31049 bp in length
* 105704 105803: gap of 100 bp
* 105804 151961: contig of 46158 bp in length.
Location/Qualifiers
1. 151961
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-77F17"
/clone_lib="RP11-11 Human Male BAC"
1. 1451
/note="assembly_fragment"
1552. 3695
/note="assembly_fragment"
3796. 6752
/note="assembly_fragment"
clone_end:T7
vector_side:right
6853. 18026
/note="assembly_fragment"
18127. 30673
/note="assembly_fragment"
30774. 44941
/note="assembly_fragment"
45042. 58179
/note="assembly_fragment"
clone_end:SP6
vector_side:left
58280. 74554
/note="assembly_fragment"
74655. 105703
/note="assembly_fragment"
105804. 151961
/note="assembly_fragment"

BASE COUNT 47790 a 28505 c 28546 g 46219 t 901 others
ORIGIN

Query Match 11.6%; Score 147.6; DB 2; Length 151961;
Best Local Similarity 73.7%; Pred. No. 3.2e-20;
Matches 202; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

QY 831 CTGCAAGACCTGCATAATCCGGTGGAGAGAGCGCCTATGTGTGTGTGTGCAGGCAACAG 890
Db 41665 CTGCAAGACCTGTGACATCCATGGGAGGCGTGTATGTGTGTGTGTATAGACCTAA 41606
QY 891 TAAAGTGTACTTCAAAACAGTCTGCCGAGTGTGTGAGAAATCCTACAAACCTTACAGAG 950
Db 41605 TAAAGTGTACTTCAAA--AGTTTACAGAACTTGTGAGAAAGTCTTATTAACCTGTAACAAG 41549
QY 951 TGGAGACATCACTGTCAAAAGTGTAAAGAACTAGATGTGCTGCCAGTGAATTC 1010
Db 41548 GGGAGATTAACAGCAATCAAAAGTGTAAATTAACCTAGATGTGCTGCCAGATTTTC 41489
QY 1011 GCCACGTGACCCCTAAAGCGCCCATCGCAAGACCTTGTGGAGATGCAAGCAAAAC 1070
Db 41488 TCTACATTAACCCCTAAAGCGCCCTATGTCAAGATTTGTGGAGATGCAAGCAAAAC 41429
QY 1071 GCCTGTCTGCGACAGCACCCTTCACCTTCAATA 1104
Db 41428 ACCTATCTGTGTGAGCAGCCTTCACTTCAATA 41395

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL	2 (bases 1 to 187578) Wang, C., Meyer, R. and Dignan, G. The sequence of Homo sapiens BAC clone RP11-191J2 Unpublished (2001) 3 (bases 1 to 187578) Waterston, R.H. Direct Submission Submitted (03-OCT-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA 4 (bases 1 to 187578) Waterston, R.H. Direct Submission Submitted (29-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA 5 (bases 1 to 187578) Waterston, R.H. Direct Submission Submitted (05-APR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA 6 (bases 1 to 187578) Waterston, R. Direct Submission Submitted (20-APR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA On Apr 5, 2002 this sequence version replaced gi:19807902.
--	---

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics -----
Center project name: H_NH0191J02

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPcrl-11 human BAC library was made from the blood of one male donor, as described by Oseegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanesse, J.J. and de Jong, P.J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-473N2; the clone sequenced to the right is RP11-757K22, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-191J2; actual end is at base position 17325 of RP11-757K22.

polymorphisms have been identified between AC032000, AC024191, and

FEATURES	source
AC096952. Data from AC032000 and AC024191 were used to finish this clone, AC096952.	Location/Qualifiers
	1. .187578

```

/misc_feature /organism="Homo sapiens"
repeat_region /db_xref="taxon:9606"
repeat_region /chromosome="4"
misc_feature /map="4"
repeat_region /clone="RP11-191J2"
repeat_region /clone_lib="RPC1-11"
misc_feature 1..151
repeat_region /note="match to EST R34463 (NID:g791364) yh85c01.r1"
misc_feature 659..698
repeat_region /rpt_family="L2"
misc_feature 1017..1100
repeat_region /rpt_family="L1"
misc_feature 4161..4181
repeat_region /note="match to EST BI496001 (NID:g15335345)"
misc_feature 4163..4477
repeat_region /rpt_family="Alu"
misc_feature 4178..4188
repeat_region /note="match to EST AA749271 (NID:g2789229) ny04d05.s1"
misc_feature 4541..4567
repeat_region /rpt_family="(TTTG)n"
misc_feature 4557..4577
repeat_region /note="similar to Homo sapiens EST AI265985 (NID:g3884143)"
misc_feature qm46c01.x1"
misc_feature 5981..5992
repeat_region /note="match to EST BI493951 (NID:g15333295)"
misc_feature 8924..9119
repeat_region /rpt_family="L1"
misc_feature 9141..10080
repeat_region /rpt_family="L1"
misc_feature 10292..10618
repeat_region /rpt_family="L1"
misc_feature 10536..11223
repeat_region /note="match to EST BG506934 (NID:g13468451)"
misc_feature 10662..10970
repeat_region /rpt_family="L1"
misc_feature 11491..11819
repeat_region /rpt_family="L1"
misc_feature 11832..12135
repeat_region /rpt_family="Alu"
misc_feature 12137..12205
repeat_region /rpt_family="MER2_type"
misc_feature 12204..12819
repeat_region /rpt_family="MER2_type"
misc_feature 12830..13102
repeat_region /rpt_family="L1"
misc_feature 13109..13255
repeat_region /rpt_family="MER2_type"
misc_feature 13256..13378
repeat_region /rpt_family="L1"
misc_feature 13379..13664
repeat_region /rpt_family="Alu"
misc_feature 13665..14289
repeat_region /rpt_family="L1"
misc_feature 14301..14515
repeat_region /note="match to EST BG995513 (NID:g14399583)"
misc_feature 14357..14790
repeat_region /note="similar to Homo sapiens EST BE735137"
misc_feature (NID:g10149129)"
misc_feature 14405..14880
repeat_region /note="similar to Mus musculus EST BI407638"
misc_feature (NID:g15168561)"
misc_feature 14409..14845
repeat_region /note="similar to EST BM912100 (NID:g19362479)"
misc_feature 14441..14845
repeat_region /note="similar to EST BQ054284 (NID:g19813624)"
misc_feature 14447..14827
repeat_region /note="similar to Homo sapiens EST BE799707"
misc_feature (NID:g10220905)"
misc_feature 14447..14813
```


FEATURES	source
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA	
Location/Qualifiers	
1. .125020	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="16"	
/map="16q24.3; between D16S520 and WI-12410"	
/note="Isolated from a patient with Huntington's Disease-Like 2 (HDL2)"	
repeat_region	
complement(35581. .35746)	
/rpt_type=tandem	
/rpt_unit=ctg	
gene	
complement(<36507. .>36887)	
/gene="JPH3"	
/note="JP3"	
mRNA	
complement(<36507. .>36887)	
/gene="JPH3"	
/product="junctophilin 3"	
CDS	
complement(<36507. .36887)	
/gene="JPH3"	
/note="component of the junctional complex between plasma membrane and endoplasmic reticulum"	
/codon_start=1	
/product="junctophilin 3"	
/protein_id="AAL40941.1"	
/db_xref="GI:17646245"	
/translation="MSSSGRFNFDDGSGYCGGWEDGKAHGHCVCPPKGGGEYTGWS	
HGFVEVLGYTWPSGNTYOGTWAQGRHIGLESKQWVYKGEWTHGFKGRYGVRECAAG	
NGAKYEGTWSNGLQDGYGTETYSDG"	
BASE COUNT	
29056 a 32731 c 30696 g 28283 t 4254 others	
ORIGIN	

Query Match	5.7%	Score 72.2	DB 9	Length 125020
Best Local Similarity	11.1%	Pred. No. 0.00011		
Matches 100	Conservative 401	Mismatches 395	Indels 4	Gaps 3
OY 4	GGGGGCGAGCGCGGAGCAGCCCATGTTCCCGGCGAGCAGTTCACCCCTGCCCCGAT	63		
Db 17783	GSKSSRGCTGYMKKKGGGKSGMSKKWKGSSTSRRRSSAKSCSYMMGSMCMSCMSMA	17724		
OY 64	CCTTATCCGAGAGCCACCAAGCCGGGATGCGCTGAGGTTGAGCCAGGGGCTGCCGA	123		
Db 17723	KSYMCMCYMYRMSSYSMSYKYKCYSCMGMSSTYSYSCCMKMSWSGSCYKCMKYSGWSS	17664		
OY 124	CCCCGCCCCCCTCCTTCCTCCCGCGGTACAGACAGCTATGACCCGCGAGTACGTGA-	182		
Db 17663	YSTRSMGYSSSTCKKTKCSMSMYKCKTSKSYRRKRSYVWGGRRAKAKKYYCAGRRRRM	17604		
OY 183	-CAGCCACCAAGCGGGGACACAGTTCATGGCCCTGCTGTCCGGATGGGTCCCCGGTCCGTCA	241		
Db 17603	SYMKCCAKWMSYCCWMSYCMITYYSKSCTYKSSCTCYKRGYYWGSKTCYSAGKSRSMY	17544		
OY 242	GCAGCCGTGACGCTGCGGTGCAG-CTGAACCCGCGCGCCGACGCTCGGTGCAGTGTCA	300		
Db 17543	CMMRSSKSSWSMSMARSSWCMGWAGYRRSKRSAGWAGAGWRSKGRSTGMKRACSKKT	17484		
OY 301	CTCGGCGCGCGACGCTGCAGCCTGCAGGGGTGCCGAGCCAGCCCCGACGCCGATCGGT	360		
Db 17483	GSYGTGRSMKKKKKYSKYSRGMKKKKTCTCYCMWKYYKYRKTSMCWYYMKSWGKYRYKR	17424		
OY 361	TCCTGTCAACCCCGTGGCCACAGCGCGCGCGGAGATCCCCGCG-ATCCTGGCAGACCGT	419		
Db 17423	CCMKKCKCTGYRGMSKKSGYKSMGMRGSSYSTSCWKS CWGYSWMMCKMKYMSYKKRRS	17364		
OY 420	AGCCCCGTTCTCGTCCGTGACCTTCTGTGGCCTCTCCTCCTACATGGAGGTTGCGGAGG	479		
Db 17363	MRGSMSSKGWRYAGRGCTSSSMSTRRRSKCYSYKKGRKMGWGMKRGSKYWS	17304		
OY 480	CAGGCACACACCCACGAAAGGAGAGGGAGACCCCGCATCTCTCGGGACCCGGGAACCGGA	539		
Db 17303	SMKKMRSSWSKCYSTKSYKSGRRRSKGWRSTKSKAKSSMRMAGSKCTYGSYSWSNNRR	17244		
OY 540	GCCCGAGAGGTTGGCCGCGAGGAAGCGGTTCCCCCAGCCGCGGAAGCGAGGCGGATGT	599		

Db	17243	NNRMGKTCNMYMRSRAMMNGNAAGCTTCCCCANTNGGGGAAAAAGGCGSASRAS	17184
QY	600	TCAGGCTGCAGGCGCAGCGCCGGGTGGAGACAGCCACCACCGAGGAGACCGGAACAGTGT	659
Db	17183	CYKGRMSSKSCYRSGTTRRCMKSKCRGRGSKGSMGMTGRGSGCKTYSAKGSGRGCIYCW	17124
QY	660	GCGGCGGATGCAGTCTGAGCCTGGAGAGCAGCCATGTCTCTCCCGCAGAGATTGGCTCA	719
Db	17123	GWGRKGRKCMSSRKMKMYKSYRRKRWMTCMKMCMYSMYMAMYCRSMCCMCMCKSCCGCY	17064
QY	720	GGACCCCGGTGATTCGATGCCCCCTCGAGACAGGCCCTCCCCGAAAGCACGGAGACGA	779
Db	17063	MGMSSYSYSGKYSWGMKMSYMMWSYISKRSTSKAMRSSKRGMTGGRHKGGGRSYGKGGS	17004
QY	780	CAAGGAGCGCCCTGCGTTTCCAGTCTCTTAGACAGAGTACGGCTACTATCACTGCAAGA	839
Db	17003	YKGGGSWGGKGGKSTRGSSAGKKSSYKKMCCCARYKSMYSKYSRMMWCMMSYIYCWGK	16944
QY	840	CTGCAAAATCCGGTGGAGAGCGCCTATGTGTGTGTGTGACAGGCGCACCAAGTAAGTGT	899
Db	16943	CWSSCWSSMRKKSKGGSWAKGMVWDKGSVSTDKSDKMBRSBSKYSKWSMSRGRKYY	16884

Search completed: January 19, 2003, 22:13:21
Job time : 5277 secs

: CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/935,313
 : FILING DATE:
 : APPLICATION NUMBER: EP 91 114 300.6
 : FILING DATE: 26-AUG-1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: BENT, Stephen A.
 : REGISTRATION NUMBER: 29,768
 : REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (703)836-9300
 : TELEFAX: (703)683-4109
 : TELEX: 899149
 : INFORMATION FOR SEQ ID NO: 14:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 7218 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : CLONE: PT29pt-F15
 :
 : US-08-232-463-14

```

Query Match      4.28;   Score 53.2;   DB 1;   Length 7218;
Best Local Similarity 6.5%;   Pred. No. 0.00085;
Matches 25; Conservative 202; Mismatches 155; Indels 0; Gaps 0;

```

OY	465	GGAGGTTGCCGGAGGCACGACACCACGAAGGAGAGGGAGCCCGGCATCTCGGG	524
Db	1388	RRR	1329
OY	525	GACCCGGGAAACCGAGCCGAGAGGTGGCCGCGAGGAACGGTCCCCAGCGGAAG	584
Db	1328	RRR	1269
OY	585	CGAGGAGGGCGATGTTCAAGCTCAGGGCAGGCGCGGTGGAGCAGCACACCAGGA	644
Db	1268	RRR	1209
OY	645	GGACCGGAACAGTGTCGGCGCATGCAGTCTGAGCCTGGGAGCAGAGCCATGCTCGC	704
Db	1208	RRR	1149
OY	705	CGCAGAGATGGCTCAGSACCCCGGTGATTGCGATGCCCTCGAGACCAGGCCCTCCCGCA	764
Db	1148	RRR	1089
OY	765	AAGCACGAGCAGACAAGAAGCGCCTGCCCTTCCAGTCTTAGAGCAGAAGTACGGCTA	824
Db	1088	RRR	1029
OY	825	CTATCACTGCAAGACTGCAA 846	
Db	1028	AATTCGTGAGCGTATGGCAA 1007	

RESULT 4
US-08-456-200B-11
; Sequence 11, Application US/08456200B
; Patent No. 6229000
; GENERAL INFORMATION:
; APPLICANT: Franz, Jurgен; Weingartner, Bernhard;
; APPLICANT: Unterbeck, Axel; Rae, Peter
; TITLE OF INVENTION: TISSUE-SPECIFIC HUMAN NEURONAL
; TITLE OF INVENTION: CALCIUM CHANNEL SUB-TYPES AND
; TITLE OF INVENTION: THEIR USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York

```

: COUNTRY: U.S.A.
: ZIP: 10591-5144
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB
: MEDIUM TYPE: storage
: COMPUTER: NEC Powermate SX/20
: OPERATING SYSTEM: DOS
: SOFTWARE: Wordperfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/456,200B
: FILING DATE: 31-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/094,712
: FILING DATE: 19-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/858,278
: FILING DATE: 26-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/064,778
: FILING DATE: 19-MAY-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE 41 10 785
: FILING DATE: 04-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Kurt G. Briscoe
: REGISTRATION NUMBER: 33,141
: REFERENCE/DOCKET NUMBER: Bayer 8398.3-KGB
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (914) 332-1700
: TELEFAX: (914) 332-1844
: TELEX:
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6232 nucleotides
: TYPE: Nucleotide
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: cDNA
: US-08-456-200B-11

```

Query Match	4.0%;	Score 50.8;	DB 4;	Length 6232;
Best Local Similarity	48.0%;	Pred. No. 0.0032;		
Matches 201; Conservative	0;	Mismatches 215;	Indels 3;	Gaps 2;

QY	250	GACGCTGCGGTGCAGGTGAACCCGCGCGGACGCGCTCGGTGCAGTGTTCACCTCGGGCGC	309
Db	1636	GAGGCCCCGAAGGGCGAGAGCGGGGAGGCCCGGTGCCCCGGAGAGCGGCCGCGCCAC	1695
QY	310	CGCAGCGCTGCAGCCTGCAGGCTGCCGAGCCAGCCAGCCCCGACGCCCGATCGGGTTCGTCAA	369
Db	1696	CGCA-GCCACAGCAAGAGAGGCGCGCGCGCC--CCCGAGCGCGGAGCAGCGCGCGCGA	1752
QY	370	CCCCGTGGCCACGCGCGCGCGGGAGATCCCCCGCATCTGCGAGACCGTAGCCCGTTC	429
Db	1753	GGCCCAAGCCCCGAGGCGCGCGCGGCGGACACCGCGCGGCTCCCGGAGAGCGCGCC	1812
QY	430	TCGTCCGTGACCTTCTGTGGCCTCTCCTCTCACTGAGCTTGCGGGAGCGACGACACA	489
Db	1813	GAGCGGAGCCCCGACGCCACCGCGGCACCGGACACGAGATCCGAGCAAGAGATGCCGC	1872
QY	490	CCCACGAAGGAGAGGCGGAGCCCGGCATCCTCGGGACCCGGGAACCGGAGCCGAGAGAG	549
Db	1873	GCGCGCAAGGGCGAGCGGNGCGCGGCACCGCGGGCGCCCCCGAGCGGGCCCCGGGAG	1932
QY	550	GTGGCCCGGAGAAAGCGGTCCCCCAGCCGCGGAAGCGAGAGAGGGCGATGTTCAAGCTTCA	609
Db	1933	GCGGAGAGCGGGGAGGAGCCGGCGCGCGCGCACCGCGCCCCGACAAAGGCGCAGCTTCT	1992
QY	610	GGGACGCGCGGTGGGAGCAGCAGCCACCAACCGAGGAGCCGGAACAGTAGTGGCGCGCAT	669
Db	1993	CACGAGGCTGTGAGAAAGAGACCAACNGAGAAGGAGGCCACCGAGAGAGGAGGCTTACAT	2051


```

: APPLICANT: Williams, Mark
: APPLICANT: Feldman, Daniel
: APPLICANT: McCue, Ann
: APPLICANT: Brenner, Robert
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: TITLE OF INVENTION: METHODS
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-2926
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,543A
: FILING DATE: May 31, 1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/223,305
: FILING DATE: April 4, 1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/868,354
: FILING DATE: April 10, 1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-AUG-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/620,250
: FILING DATE: 30-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/482,384
: FILING DATE: 20-FEB-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/603,751
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US89/01408
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/176,899
: FILING DATE: 04-APR-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-52517
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619)238-0999
: TELEFAX: (619)238-0062
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7175 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 144..6857
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..143
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 6855..7175
:
: US-08-455-543A-8

```

	Matches	201; Conservative	0; Mismatches	215; Indels	3; Gaps	2;
QY	250	GACGCTGCGGTGCAGGTGAACCCCGCCGCGACGCCCTCGGTGCAGTGTTCACCTCGGGCGC				
Db	2751	GAGGCCCGGAAGGGCGAGAGCGGGGAGGCCCGGTGCCCGGAGAGCGGCCCGCGCCGAC				
QY	310	CGCAGCGCTGCAGCCTGCAGGGTGGCCGAGCCAGGCCCGACGCCCGATCGGGTTCGTCAA				
Db	2811	CGCA-GCCACAGCAAGAGAGCGCGGGGCC--CCCGAGGCGCGGAGCGAGCGCGGCCGA				
QY	370	CCCCGTGGCCACGCGCGCGCGGGAGATCCCCCGCATCCTGGCAGACCGTAGCCCCGTTTC				
Db	2868	GCCCCAGGCCCGGAGGGCGCGCGCGGCACACACCGCGCGGCTTCCCCGGAGAGGCGGCC				
QY	430	TCGTCCGTGACCTTCTGTGGCCTCTCCTCTCACTGGAGTTTCGGGAGGCGAGGCAGACA				
Db	2928	GAGCGGGAGCCCCGACGCCACCGCGCGGCACCGGCACCAAGATCCGAGCAAGGAGTGGCC				
QY	490	CCCACGGAAGGAGAGGGGAGCCCCGGCATCTCGGGGAGCCCGGGAACCGGACCGAGAGAG				
Db	2988	GCGGCCAAGGGCGAGCGGGCGCGCGGCACCGCGCGGCCCGCCGAGCGGGCCCCGGGAG				
QY	550	GTGGCCGCGAGGAAAGCGGTCCCCAGCGCGGAAAGCGAGGAGGGCGATGTTACGGCTGCA				
Db	3048	GCGGAGAGCGGGGAGGAGCGCGCGCGCGCGGCACCGGGCCCGGCACAAGGCGCAGCCTGCT				
QY	610	GGGCAGGCGGGGTGGAGCAGCACCCACCGGAGGAGACCGGAACAGTGTGGCGCGCAT				
Db	3108	CACGAGGCTGTGGAGAGAGCAGCACCGGAGAGGAGGCCACCGAAGGAGGCTGAGAT				

RESULT 8
 US-08-193-078B-8
 ; Sequence 8, Application US/08193078B
 ; Patent No. 5846757
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Harpold, Michael
 ; APPLICANT: Ellis, Steven
 ; APPLICANT: Williams, Mark
 ; APPLICANT: Feldman, Daniel
 ; APPLICANT: McCue, Ann
 ; APPLICANT: Brenner, Robert
 ; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 ; TITLE OF INVENTION: METHODS
 ; NUMBER OF SEQUENCES: 29
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
 ; STREET: 1660 UNION STREET
 ; CITY: SAN DIEGO
 ; STATE: CA
 ;
 ; COUNTRY: USA
 ;
 ; ZIP: 92101
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/193,078B
 ; FILING DATE: 07-FEB-1994
 ;
 ; CLASSIFICATION: 435
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/868,354
 ; FILING DATE: 10-APR-1992
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/745,206
 ; FILING DATE: 15-AUG-1991
 ;
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L.
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 6362-53607
 ;
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-238-0999

	ADDRESSSEE:	Brown, Martin, Haller & McClain
	STREET:	1660 Union Street
	CITY:	San Diego
	STATE:	California
	COUNTRY:	US
	ZIP:	92101
	COMPUTER READABLE FORM:	
	MEDIUM TYPE:	Floppy disk
	COMPUTER:	IBM PC compatible
	OPERATING SYSTEM:	PC-DOS/MS-DOS
	SOFTWARE:	Patentln Release #1.0, Version #1.25
	CURRENT APPLICATION DATA:	
	APPLICATION NUMBER:	US/08/949,386
	FILING DATE:	
	CLASSIFICATION:	
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER:	US/08/290,012
	FILING DATE:	11-AUG-1994
	APPLICATION NUMBER:	08/149,097
	FILING DATE:	5-NOV-1993
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER:	08/105,536
	FILING DATE:	11-AUG-1993
	ATTORNEY/AGENT INFORMATION:	
	NAME:	Seldman, Stephanie L.
	REGISTRATION NUMBER:	33,779
	REFERENCE/DOCKET NUMBER:	519808
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE:	(619) 238-0999
	TELEFAX:	(619) 238-0062
	INFORMATION FOR SEQ ID NO:	8
	SEQUENCE CHARACTERISTICS:	
	LENGTH:	7175 base pairs
	TYPE:	nucleic acid
	STRANDEDNESS:	double
	TOPOLOGY:	linear
	MOLECULE TYPE:	DNA (genomic)
	FEATURE:	
	NAME/KEY:	CDS
	LOCATION:	144..6857
	FEATURE:	
	NAME/KEY:	5'UTR
	LOCATION:	1..143
	FEATURE:	
	NAME/KEY:	3'UTR
	LOCATION:	6855..7175
	US-08-949-386-8	
	Query Match	3.8%; Score 49; DB 3; Length 7175;
	Best Local Similarity	48.0%; Pred. No. 0.0094;
	Matches 201; Conservative	0; Mismatches 215; Indels 3; Gaps 2;
QY	250 GACGCTGGCGTGAGGTGAACCCGCCGCCGACGCCTCGTGTGACATGTTCACTCGGGCGC	309
Db	2751 GAGGCCCCGAAGGGCGAGAGCGGGGAGCCCGGTGCCCCGGAGAGAGCGCGCGCCGAC	2810
QY	310 CGCACGCTGCAGCCTTCGACGGGTGCCGAGCCAGCCCCGACGCCCGATCGGTTCTCTCAA	369
Db	2811 CGCA-GCCACAGCAAGAAGAGCGCCCGCGGCC--CCCGAGGCGCGAGCGCGCGCGA	2867
QY	370 CCCCCTGGCCACGCGCGCGCGCGAGATCCCCCGCATCTCTGGCAGACCGTAGCCCCGTT	429
Db	2868 GGCCCAAGGCCCGGAGGGCGCGCGCGCACACCAGCGGCTCCCGGAGGAGGCGGCC	2927
QY	430 TCGTCCGTGACCTTCTGTGCTCTCTCTCTCTCACTGGAGTTGCGGAGGAGGACAGACA	489
Db	2928 GAGCGGGAGCCCCGACGCCACCGCGCGCACCGGACACAGATCCGAGCAAGAGTGCGCC	2987
QY	490.CCCACGAAGGAGAGAGGGAGCCCGCATCTCTGGGGACCCCGGAACCGAGCCGAGAGAG	549
Db	2988 GGCGCCAAAGGGCGAGCGCGCGCGCGCACCGCGCGGCGGCCCCCGAGCGGGCGGAG	3047
QY	550 GTGGCTCGGAGGAAGCGGTCCCCCAGCCCGCAAGCGAGGAGGCGCATGTTCAAGCTCA	609

```

Db      3048  CCGGAGAGCGGGGAGAGAGCCGGCGCGCGGACACCGGGCCGGCACAAGGCGAGCTGCT 3107
QY      610  GGGCAGGCCGGGTGGAGCAGCACCGCACCGGAGAGGAGCGGAGACAGTGTGGCGGCGAT 668
Db      3108  CACGAGGCTGTGGAGAGAGGAGCACCAGGAGAGGAGGCCACAGGAGAGAAGGAGGCTGAGAT 3166

RESULT 12
US-08-450-562-8
; Sequence 8, Application US/08450562
; Patent No. 6096514
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Gillespie, Alison
; APPLICANT: Feldman, Daniel
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: US
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,562
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/404,950
; FILING DATE: 13-MAR-1995
; APPLICATION NUMBER: 08/336,257
; FILING DATE: 7-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/314,083
; FILING DATE: 28-SEPT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/311,363
; FILING DATE: 23-SEPT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290,012
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/223,305
; FILING DATE: 4-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/193,078
; FILING DATE: 07-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,097
; FILING DATE: 5-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105,536
; FILING DATE: 11-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/914,231
; FILING DATE: 13-JULY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: 10-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06903

```

FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/603,751
FILING DATE: 08-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/482,384
FILING DATE: 02-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-519812
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-450-562-8

Query Match 3.8%; Score 49; DB 3; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0094;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGTGCAGGTGAACCCGCGGAGCGCTCGGTGCACTGTTCCGCGGC 309
DB 2751 GAGGCCCCGAGAGGCGAGAGCGGAGAGCCCGTGCCTCCGCGGAGAGAGCGCGCCGCGCAC 2810
QY 310 CGCAGCTGCAGCTGCAGGGTGCCGAGCCAGCCCGGATCGGGTTCCTGTCAA 369
DB 2811 CGCA-GCCACAGCAAGAGCGCGCGGGCC--CCCGAGGCGCGGAGCGCGCGCGCA 2867
QY 370 CCCCCTGGCCACGCCCGCGGGAGATCCCCCGCATCTGGCAGACCGTAGCCCGCTTC 429
DB 2868 GCGCCAGGCCCCGAGGCGCGCGCGCGGACACCGCGCGCTCCCGGAGAGCGCGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGAGCGAGACA 489
DB 2928 GAGCGGGAGCCCCGACGCCACCGCGCGCAGCCGCGATCCGAGCAAGAGATGCGCC 2987
QY 490 CCCACGAGGAGAGGAGCGCGCGCGCATCTCGGGGAGCCCGGAGCGAGCGAGAGAG 549
DB 2988 GCGCGCAAGGCGAGCGCGCGCGCGGACCGCGCGCGCGCGCGAGCGCGCGCGGAG 3047
QY 550 GTGGCCGCGAGAAACCGCTCCCGCAGCGCGGAGAGCGAGGCGATTTCAAGCTGCA 609
DB 3048 GCGGAGAGCGGAGAGCGCGCGCGCGCGGACCGCGCGCGCGCGCGAGCGCGCTGCT 3107
QY 610 GCGCAGGCGCGGTGGAGCAGCAGCCACCGAGAGAGCGGAGACAGTGTGGCGCGCAT 668

DB 3108 CACGAGCTGTGAGAGAGAGACCACCGAGAGAGGCGCACGAGAGAGCGCTGACAT 3166
RESULT 13
US-08-984-709A-8
Sequence 8, Application US/08984709A
Patent No. 6320032
GENERAL INFORMATION:
APPLICANT: Williams, Mark E.
APPLICANT: Stauderman, Kenneth A.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, Suite 700
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,709A
FILING DATE: 02-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 450-8400
TELEFAX: (619) 587-5360
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-984-709A-8

Query Match 3.8%; Score 49; DB 4; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0094;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
QY 250 GACGCTGCGTGCAGGTGAACCCGCGGAGCGCTCGGTGCACTGTTCCGCGGC 309
DB 2751 GAGGCCCCGAGAGGCGAGAGCGGAGAGCCCGTGCCTCCGCGGAGAGAGCGCGCCGCGCAC 2810
QY 310 CGCAGCTGCAGCTGCAGGGTGCCGAGCCAGCCCGGATCGGGTTCCTGTCAA 369
DB 2811 CGCA-GCCACAGCAAGAGAGCGCGCGGGCC--CCCGAGGCGCGGAGCGAGCGCGCGCA 2867
QY 370 CCCCCTGGCCACGCCCGCGGGAGATCCCCCGCATCTGGCAGACCGTAGCCCGCTTC 429
DB 2868 GCGCCAGGCCCCGAGGCGCGCGCGCGGACACCGCGCGCGCTCCCGGAGAGCGCGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTTGCGGAGAGCGAGACA 489

Db 2928 GAGCGGAGCCCCGACGCCACCGCGCCACCGGACACGATCCGAGCAAGAGTCCGCC 2987
QY 490 CCCACGAGGAGAGGGGAGCCCCGCGATCTCGGGAGCCCCGGAACCGAGCCGAGAGAG 549
Db 2988 GCGGCCAAGGGCGAGCG 3047
QY 550 GTGGCCCGGAGAAAGCGGTCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
Db 3048 GCGGAGAGCGGGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3107
QY 610 GGGCAGCGCGGTGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 668
Db 3108 CACGAGCGTGTGAGAAAGAGAGACGACGACGAGAGAGAGGCGCACGAGAGAGGCTGAGAT 3166

RESULT 14

US-08-450-272-8
Sequence 8, Application US/08450272
Patent No. 6387696

GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Alison
APPLICANT: Feldman, Daniel
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101

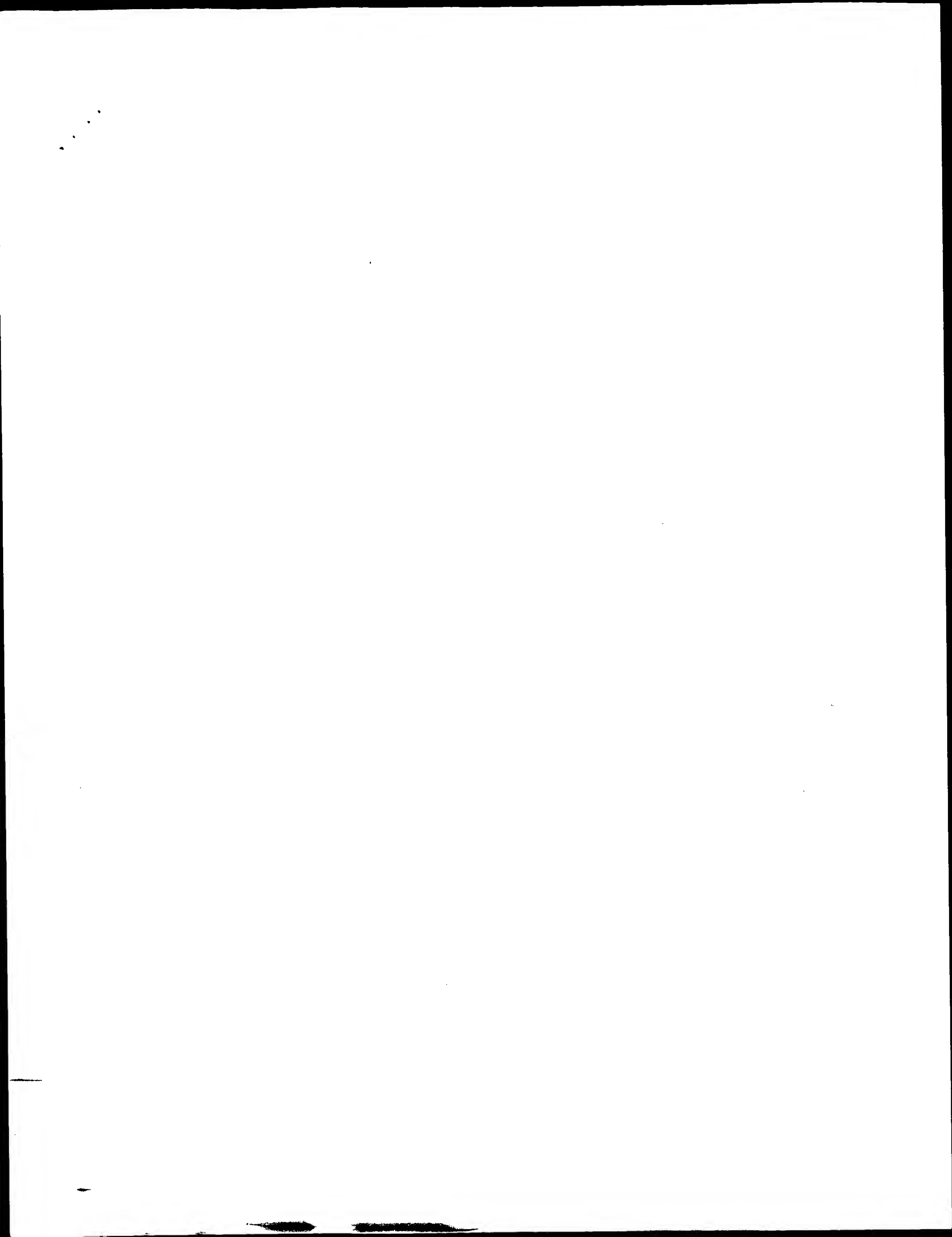
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,272
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,950
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: 08/336,257
FILING DATE: 7-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/314,083
FILING DATE: 28-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/311,363
FILING DATE: 23-SEPT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,012
FILING DATE: 11-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: 4-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/193,078
FILING DATE: 07-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,097

FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06903
FILING DATE: 14-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-519812
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-450-272-8

Query Match 3.88; Score 49; DB 4; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.0094;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGGTGAGTGAACCCGCGCGGACGCTCGGTGAGTGTTCATCTGGGGCG 309
Db 2751 GAGGCCCCGAAGGCGGAGAGCGGGAGCCCGGTGCGCGGAGAGAGCGCGCGCGCAC 2810
QY 310 CGCAGCTGACGCTGACAGGTGCGGAGCCGAGCCGCGGATCGGATCTCTGTCAA 369
Db 2811 CGCA-GCCACAGCAAGAGAGCGCGGGGCC--CCCGAAGCGCGAGCGAGCGCGCGCA 2867
QY 370 CCCCCTGGCCACGCGCGCGCGGAGATCCCCCGGATCTTGCGACGACCTAGCCCCGTT 429
Db 2868 GGGCCAGGCCCCGAGGCGCGCGCGCGGACACCGCGCGCGCTCCCCGAGAGCGCGGCC 2927
QY 430 TCGTCCGTGACCTTCTGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
Db 2928 GAGCGGAGAGCCCGACGCGCGCGCGCGGACCGGACCGGACCGGATCCGAGCAAGAGTGGCC 2987
QY 490 CCCACGAGGAGAGAGGAGCGCGCGCATCTCGGGGAGCCCGGAACCGGAGCGAGAGAG 549
Db 2988 GCGGCCAAGGGCGAGCGCGCGCGCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGGAG 3047
QY 550 GTGGCCCGGAGAAAGCGTCCCCAGCGCGGAGCGAGCGAGGCGGATGTTAGGCTGCA 609
Db 3048 GCGGAGAGCGGGGAGAGAGCGCGCGCGCGCGCGGACCGCGCGCGCGCGCACAGGCGGCTGCT 3107



QY	1	AAGCGGGGCGAGGCCGGACGCACCATTGTTCCTCCGGCAGACACTTCCACCCCCTGCCCC	60
Db	1	AAGCGGGGCGAGGCCGGACGCACCATTGTTCCTCCGGCAGACACTTCCACCCCCTGCCCC	60
QY	61	CATCCTTTATCCGGCAGGCCACCACAAGCCGGGGATGGCTGGAGGTTCCGAGCCAGGGGCTGC	120
Db	61	CATCCTTTATCCGGCAGGCCACCACAAGCCGGGGATGGCTGGAGGTTCCGAGCCAGGGGCTGC	120
QY	121	CGACCCGGCGCCCCCTCTCTTCCTCCCGGGCTACAGACAGCTCATGGCCGGGAGTAGCTTC	180
Db	121	CGACCCGGCGCCCCCTCTCTTCCTCCCGGGCTACAGACAGCTCATGGCCGGGAGTAGCTTC	180
QY	181	GACAGCCACCAGCGGGCACAGCTATGGCCCTGCTGTCCGGATGGGTCCCCGGTGCGTTC	240
Db	181	GACAGCCACCAGCGGGCACAGCTATGGCCCTGCTGTCCGGATGGGTCCCCGGTGCGTTC	240

Db	181	GACAGCCACCAGCGGGCACAGCTCATGTGGCCCTGCTGTGCGGATGGGTCCCGGTGCGT	240
QY	241	AGCAGCCGTGACGCTGCGGTGCAGGTGAACCCCGCCGACGCCCTCGGTGCAGTGTTC	300
Db	241	AGCAGCCGTGACGCTGCGGTGCAGGTGAACCCCGCCGACGCCCTCGGTGCAGTGTTC	300
QY	301	CTCGGGCGCCGACGCTGCAGCCTGCAGGGTGTCCGAGCCAGCCCGACGCCGATCGGGT	360
Db	301	CTCGGGCGCCGACGCTGCAGCCTGCAGGGTGTCCGAGCCAGCCCGACGCCGATCGGGT	360
QY	361	TCCTGTCAACCCCGTGGCCACGCGCGCGCGGAGATCCCCCGGATCCTGGCAGACCCTA	420
Db	361	TCCTGTCAACCCCGTGGCCACGCGCGCGGAGATCCCCCGGATCCTGGCAGACCCTA	420
QY	421	GCCCCGTTCTCGTCCGTGACCTTCTGTGGCCCTCTCCTCACTGGAGGTTGCGGAGGC	480
Db	421	GCCCCGTTCTCGTCCGTGACCTTCTGTGGCCCTCTCCTCACTGGAGGTTGCGGAGGC	480
QY	481	AGGCAGACACCCACGAGGGAGAGGGAGCCCGGCATCCTGGGGACCCGGGAACCGGAG	540
Db	481	AGGCAGACACCCACGAGGGAGAGGGAGCCCGGCATCCTGGGGACCCGGGAACCGGAG	540
QY	541	CCGAGAGAGGTGGCCGCGAGGAAAGCGTCCCCAGCCCGGAAGCGAGAGGCGCATGTT	600
Db	541	CCGAGAGAGGTGGCCGCGAGGAAAGCGTCCCCAGCCCGGAAGCGAGAGGCGCATGTT	600
QY	601	CAGGCTGCAGGGCAGGCGCGGTGGGAGCAGCACCACCGGAGACCGGAACAGTGTG	660
Db	601	CAGGCTGCAGGGCAGGCGCGGTGGGAGCAGCACCACCGGAGACCGGAACAGTGTG	660
QY	661	GCGGCGATGCAGTCTGAGCCTGGGAGCGAGAGCCATGTCTTCCCGCAGAGATGGCTCAG	720
Db	661	GCGGCGATGCAGTCTGAGCCTGGGAGCGAGAGCCATGTCTTCCCGCAGAGATGGCTCAG	720
QY	721	GACCCCGGTGATTCGATGCCCCCTCGAGACCAGGCCCTCCCGCAAAAGCACGAGCAGGAC	780
Db	721	GACCCCGGTGATTCGATGCCCCCTCGAGACCAGGCCCTCCCGCAAAAGCACGAGCAGGAC	780
QY	781	AAGGAGCGCCTGCGTTTCCAGTTCTTAGAGCAGAAAGTGGCTACTATACCTGCAGGAC	840
Db	781	AAGGAGCGCCTGCGTTTCCAGTTCTTAGAGCAGAAAGTGGCTACTATACCTGCAGGAC	840
QY	841	TGCAAAATCCGTTGGGAGAGCGCCTATGTGTGTGTGTCAGGGCACCAGTAAGGTGTTA	900
Db	841	TGCAAAATCCGTTGGGAGAGCGCCTATGTGTGTGTGTCAGGGCACCAGTAAGGTGTTA	900
QY	901	CTTCAACAGTCTGCGCAGTGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGACAT	960
Db	901	CTTCAACAGTCTGCGCAGTGTGTGAGAAATCCTACACCCCTTACAGAGTGGAGACAT	960
QY	961	CACCTGTCAAAAGTTGTAAAGAAGTATGATGTGCTGCCAGTCAGATTTGGCCACGTGA	1020
Db	961	CACCTGTCAAAAGTTGTAAAGAAGTATGATGTGCTGCCAGTCAGATTTGGCCACGTGA	1020
QY	1021	CCCTAAACGCCCCCATGGCAAGACTTGTGTGGAGATGCAAAGCAGCCTGTCTG	1080
Db	1021	CCCTAAACGCCCCCATGGCAAGACTTGTGTGGAGATGCAAAGCAGCAGCCTGTCTG	1080
QY	1081	CGACAGCACCTTCAGCTTCCAAATACATCATTTAGTGAAGTGCAGAAACGTTTCTGTAGA	1140
Db	1081	CGACAGCACCTTCAGCTTCCAAATACATCATTTAGTGAAGTGCAGAAACGTTTCTGTAGA	1140
QY	1141	TGGGGCTAATGGAATGACAAGTGAAGCTTCTCCCTCTTCACTCTCCCTTCCAAAT	1200
Db	1141	TGGGGCTAATGGAATGACAAGTGAAGCTTCTCCCTCTTCACTCTCCCTTCCAAAT	1200
QY	1201	TCTTCATGACAGACAGTGTACTTGGATATAAAGCCTGTGAATAAAAGTATTGCAACA	1260
Db	1201	TCTTCATGACAGACAGTGTACTTGGATATAAAGCCTGTGAATAAAAGTATTGCAACA	1260
QY	1261	AAAAAAAAAAAAAAAA 1277	
Db	1261	AAAAAAAAAAAAAAAA 1277	

```

RESULT 2
US-09-844-864-18
; Sequence 18, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 814
; TYPE: DNA
; ORGANISM: mus musculus
US-09-844-864-18

```

Query Match	62.6%;	Score 799.4;	DB 10;	Length 814;
Best Local Similarity	99.9%;	Pred. No. 4.7e-189;		
Matches 800;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0

QY	3	GGCGGGCGAGCGCGGGGACGGACCCATTGTTCCCGGGCAGACAGTTCCACCCCTGCCCGCA	62
Db	1	GGCGGGCGAGCGCGGGGACGGACCCATTGTTCCCGGGCAGACAGTTCCACCCCTGCCCGCA	60
QY	63	TCCTTATCCGACAGGCCACCAAGCCGGGGATGGCTGGAGGTTCCGAGCCAGGGGCTGCCG	122
Db	61	TCCTTATCCGACAGGCCACCAAGCCGGGGATGGCTGGAGGTTCCGAGCCAGGGGCTGCCG	120
QY	123	ACCCGGCCCCCTCCTTCTCTCCCGGCTACAGACAGCTCATGGCCCGGAGTAAGTACGTCGA	182
Db	121	ACCCGGCCCCCTCCTTCTCTCCCGGCTACAGACAGCTCATGGCCCGGAGTACGTCGA	180
QY	183	CAGCCACCAGCGGGGCACAGCTCATGGCCCTGCTGTCCGCGATGGTCCCCGGTCCGTACG	242
Db	181	CAGCCACCAGCGGGGCACAGCTCATGGCCCTGCTGTCCGCGATGGTCCCCGGTCCGTACG	240
QY	243	CAGCCGTGACCGTGGCGGTGCAGGTGAAACCCGCGCGACGCGCTCGGTGCAGTGTTCAC	302
Db	241	CAGCCGTGACCGTGGCGGTGCAGGTGAAACCCGCGCGACGCGCTCGGTGCAGTGTTCAC	300
QY	303	CGGGCGCCGACAGCGCTGCAGCCTGCAGGGTGCAGGCCAGCCCGGACCGCGATCGGGTTC	362
Db	301	CGGGCGCCGACAGCGCTGCAGCCTGCAGGGTGCAGGCCAGCCCGGACCGCGATCGGGTTC	360
QY	363	CTGTCAACCCCGTGGGCCACGCGCGCGCGGGGAGATCCCCCGCATTCCTGGCAGACCGTAAGC	422
Db	361	CTGTCAACCCCGTGGGCCACGCGCGCGCGGGGAGATCCCCCGCATTCCTGGCAGACCGTAAGC	420
QY	423	CCCGTTCGTCCGTGACCTTCTGTGGCCTCTCCTCTCTACTGTGAGGTTGCGGGAGGCAG	482
Db	421	CCCGTTCGTCCGTGACCTTCTGTGGCCTCTCCTCTCTACTGTGAGGTTGCGGGAGGCAG	480
QY	483	GCAGACACCCACGAAGGAGAGAGGGAGACCCGGCATCTCGGGGACCCGGGAACCGAGCC	542
Db	481	GCAGACACCCACGAAGGAGAGAGGGAGACCCGGCATCTCGGGGACCCGGGAACCGAGCC	540
QY	543	GAGAGAGGTGGCCGCGAGGAAGCGGTCCCCAGCCGCGAAGCGAGGAGGGGATTTCA	602
Db	541	GAGAGAGGTGGCCGCGAGGAAGCGGTCCCCAGCCGCGAAGCGAGGAGGGGATTTCA	600
QY	603	GGCTGCAGGGCAGGCGGGGTGGGAGCAGCACCACACCGGAGGACCCGGAACAGTGTGC	662
Db	601	GGCTGCAGGGCAGGCGGGGTGGGAGCAGCACCACACCGGAGGACCCGGAACAGTGTGC	660

ORGANISM: mus musculus
US-09-844-864-20

Query Match 5.6%; Score 72; DB 10; Length 105;
Best Local Similarity 100.0%; Pred. No. 5.2e-09;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 899 TACTTCAACAGTTCGCCGAGTGTGTGAGAATCTTACAACCTTACAGAGTGAGGAC 958
|||||
Db 19 TACTTCAACAGTTCGCCGAGTGTGTGAGAATCTTACAACCTTACAGAGTGAGGAC 78
QY 959 ATCACCCTGTCAA 970
|||||
Db 79 ATCACCCTGTCAA 90

RESULT 9
US-09-844-864-24

; Sequence 24, Application US/09844864

; Patent No. US20020042926A1

; GENERAL INFORMATION:

; APPLICANT: Matzuk, Martin

; APPLICANT: Ren, Yongsheng

; APPLICANT: Wu, Xuemei

; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS

; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48

; CURRENT APPLICATION NUMBER: US/09/844,864

; CURRENT FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: 60/106,020

; PRIOR FILING DATE: 1998-10-28

; PRIOR APPLICATION NUMBER: PCT/US99/25209

; PRIOR FILING DATE: 1999-10-28

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 24

; LENGTH: 105

; TYPE: DNA

; ORGANISM: mus musculus

; US-09-844-864-24

Query Match 5.5%; Score 70.4; DB 10; Length 105;
Best Local Similarity 98.6%; Pred. No. 1.3e-08;
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 899 TACTTCAACAGTTCGCCGAGTGTGTGAGAATCTTACAACCTTACAGAGTGAGGAC 958
|||||
Db 19 TACTTCAACAGTTCGCCGAGTGTGTGAGAATCTTACAACCTTACAGAGTGAGGAC 78
QY 959 ATCACCCTGTCAA 970
|||||
Db 79 GTCACCCTGTCAA 90

RESULT 10

US-09-954-456-1601

; Sequence 1601, Application US/09954456

; Patent No. US20020115057A1

; GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc

; FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456

; CURRENT FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US/60/233,617

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,052

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,134

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,637

; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1601
; LENGTH: 14800
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-1601

Query Match 3.9%; Score 49.6; DB 10; Length 14800;
Best Local Similarity 48.3%; Pred. No. 0.015;
Matches 139; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 507 GAGCCCGCATCTCTGGGAGCCCGGGAACCGGAGCCGAGAGAGTGGCCCGGAGAAAGC 566
|||||
Db 4122 GAGCCAGTACATCAAGTTCATCAGCGAGACTCTGCGCGCATGAGAGAGAGAGAGCT 4181
QY 567 GGTCCCGCAGCCGCGAGAGGAGGCGATGTTACAGCTGAGGAGCCCGGCTGGCA 626
|||
Db 4182 GGCTGAGCAGCAGCGGCGGAGAGAGAGCGGAGCGGCTGGCGAGTGGAGCGCGCTGGA 4241
QY 627 GCAGCAGCCACACCGGAGAGCGGAGAGTGTGCGCGGATGCAGTCTGAGCGCTGGAG 686
|||||
Db 4242 GAGCAGCGGCGAGCTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 4301
QY 687 CGAGGAGCCATGTCCTGCGGAGAGATGCTCAGAGACCGCGGATTCGATTCGATTCG 746
|||||
Db 4302 GAGGAGCTGCAGCAGCGGATGCAGAGAGAGAGTGTGCGGCGGAGAGAGCGGCGGCTGGA 4361
QY 747 AGACCAAGGCTCCCGCAAGCAGCAGGAGCAGGAGAGAGAGAGCGGCTGGC 794
|||
Db 4362 CGCGCAGCAGCAGAGAGCGGAGGATTCAGAGAGAGCTGCAGCAGCTGGC 4409

RESULT 11

US-10-033-026-7

; Sequence 7, Application US/10033026

; Patent No. US20020147309A1

; GENERAL INFORMATION:

; APPLICANT: Lipscombe, Diane

; APPLICANT: Schorge, Stephanie

; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF

; FILE REFERENCE: B1055/7000

; CURRENT APPLICATION NUMBER: US/10/033,026

; CURRENT FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: US 09/268,163

; PRIOR FILING DATE: 1999-03-12

; PRIOR APPLICATION NUMBER: US 60/077,901

; PRIOR FILING DATE: 1998-03-13

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 7

; LENGTH: 7177

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 146..6856

; US-10-033-026-7

Query Match 3.8%; Score 49; DB 12; Length 7177;
Best Local Similarity 48.0%; Pred. No. 0.016;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY	250	GACGCTGCGGTGACGHTGAACCCGCGCCGACGCGCTCGGTGACGTGTTCACTTCGGGCGC	309
Db	2753	GAGGCCCGAAGGCGGAGAGCGGGAGCCCGGTGCCCCGGGAGGAGCGGCCGCGCGCCAC	2812
QY	310	CGCACGCTGACGCTGCAGGGTGCAGGTCAGCCAGCCCGGACCGGATCGGGTTCCTGTCAA	369
Db	2813	CGCA-GCCACAGCAAGAGGCGCGGGGCC--CCCGAAGCGCGGAGCGAGCGCGCCGA	2869
QY	370	CCCCGTGGCCACGCGCGCGCGGAGATCCCCCGCATCTTGACAGACCGTAGCCCGTTC	429
Db	2870	GGCCCAAGCCCCGAGGCGCGCGCGGCACACCGCGCGGCTCCCGAGAGCGCGCC	2929
QY	430	TCGTCCGTGACCTTCTGTGSCCTCTCCTCTCACTGAGGTTGCGGAGCGACGACACA	489
Db	2930	GAGCGGGAGCCCCGACGCCACCGCGGCACCGCACACAGATCCGAGCAAGAGTGGCC	2989
QY	490	CCCAACGAAGGAGAGGGGAGCCCCGCATCTCGGGGACCCGGGAACCGGACCGAGAGAG	549
Db	2990	GCGGCCAAGGGCGAGCGCGCGCGGCACCGCGCGCGCCCCGAGCGGGCCCCGGGAG	3049
QY	550	GTTGCCCGCGAGGAAGCGGTTCCCCACGCCGCGGAAGCGAGGAGGGGATGTTACGGTCA	609
Db	3050	GCGGAGAGCGGGGAGGAGCCCGCGCGCGCGGCACCGGGCCCCGCGCACAAAGCGCAGCTGCT	3109
QY	610	GGGCAAGCCCGGTGGAGCAGCAGCCACACCGGAGGACCGGAAACAGTGTGGCGCGAT	668
Db	3110	CACGAGGCTGTGAGGAAGGAGCACCGGAGGAAGGAGGCCACCGAGCAAGAGGAGGCTGAGAT	3168

RESULT 12

```

US-09-954-456-1179
; Sequence 1179, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1179
; LENGTH: 7364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1179

```

Query Match	3.88;	Score 49;	DB 10;	Length 7364;
Best Local Similarity	48.08;	Pred. No: 0.016;		
Matches 201; Conservative	0;	Mismatches 215;	Indels 3;	Gaps 2;

QY	250	GACGCTGCAGTCAAGTTGAACCCGCGCCGGCAGCCCTGGTGCAGTGTTCACTCGGGCCG	309
Db	2753	GAGGCCCGAAGCGCGAGACGGGGAGCCCGGTGCCGGGAGGAGCGGCCGCGCCGAC	2812
QY	310	CGCACGCTGACGCTTCAGGCTGCCGAGCCAGCCCCGACGCCCGATCGGTCTCTGTCAA	369
Db	2813	CGCA-GCCACAGCAAGAGGCGCGGGGCC--CCGGAGGCGCGAGCCAGCCGCGCCA	2869
QY	370	CCCCGTGGCCACGCGCGCGCGGAGATCCC CGCATCTGGCAGACCCTAGCCCCGTTTC	429
Db	2870	GGCCCAAGCCCCGAGGGCGGCGCGGCACCAACCGCGGCTCCCGAGAAGCGGCC	2929
QY	430	TGCTCCGTGACCTTCTGTGGCCTCTCTCTCACTGAGGTTGCGGAGCGACAGACA	489
Db	2930	GACCGGGAGCCCCGACGCCACCGCGCACCGGCGAACAGATCCGAGCAAGAGTGC GCC	2989
QY	490	CCCACGAAGGAGAGGGGAGCCCCGGCATCTCGGGGSAACCCGGAAACGAGCCGAGAGAG	549
Db	2990	GGCGCCAAAGGGCGAGCGCGCGCGCGGCACCGCGGGGCCCCCGAGCGGGCCCCGGAGG	3049
QY	550	GTTGCCCCGAGGAAGCGTCCCCCAGCCGCGAAGCGAGAGGGCGATGTTCAAGCTGCA	609
Db	3050	GCGGAGAGCGGGGAGGAGCGCGCGCGCGGCACCGGGCCCCGACACAAGGCGCAGCTGCT	3109
QY	610	GGGCAAGCCCCGGTGGAGCAGCAACCAACCGAGAGACGGAACAGTGTGGCGGCGAT	668
Db	3110	CACGAGGCTGTGAGAGAGGAGCAACGAGAGAGAGAGGCGCAAGAGAGAGGCTGAGAT	3168

RESULT 13

```

US-10-033-026-5
; Sequence 5, Application US/10033026
; Patent No. US20020147309A1
; GENERAL INFORMATION:
; APPLICANT: Lipscombe, Diane
; APPLICANT: Schorge, Stephanie
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/10/033,026
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 09/268,163
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/077,901
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 7364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 146..7162
;
US-10-033-026-5

```

Query Match	3.88;	Score 49;	DB 12;	Length 7364;
Best Local Similarity	48.08;	Pred. No. 0.016;		
Matches 201;	Conservative 0;	Mismatches 215;	Indels 3;	Gaps 2;

QY 250 GACGCTGGCGTGCAGGTGAACCCGCCGCCGACGCTCGGTCAGTGTTCAC TC GGCGC 309
|| || || | | | | | | | | | | | | | | |
Db 2753 GAGCCCCCGAAGGGCGAGAGCGGGAGACCCCGTGCCCGGSAAGAAGCGCCGCCGCAC 2812

QY 310 CGCACGCTGCAGCCTGCAGGTTGCCGAGCCAGCCCCCGACGCCCATCGGTTCTGTCAA 369
|||| | | ||| | | | | | | | | | | | | | |
Db 2813 CGCA-GCCACAGCAAGGAGGCGCCGCCGGGCC - -CCCGAGAGCGCGGAGAGCGCGGCCGA 2869

QY 370 CCCC GTGGCCACGCGCGCGCGGAGATCCCCCGGATCCTGGCAGAACCGTAGCCCCCTTC 429
|| || || | | | | | | | | | | | | | | | |
Db 2870 GGCCCAAGGCCCGGAGGGCGCGCGCGGACCAACCGGGCGGCTCCCCCGGAGAGGGCGGCC 2929

QY 430 TCGTCGTGACCCTTCTGTGGCTCTCCTCCTCACTGAGGTTCCGGGAGGCAGGCAGACA 489

Db 2930 GAGCGGAGCCCCGACGCCACCGCCGACCGGACACCATTCGAGCAAGAGTGGCC 2989
QY 490 CCCAGAGAGGAGAGGGAGCCCGCATCTCGGGACCCGGGAACCGAGCCGAGAG 549
Db 2990 GCGGCCAAGGGGAGCGCGCCCGCCGACCCGCGCGCGCGCGCGCGCGCGAG 3049
QY 550 GTGGCCCGGAGGAAGCGGTCCCGACCGCGGAGCGAGAGGGCGATGTTCAAGCTGA 609
Db 3050 GCGGAGAGCGGGGAGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 3109
QY 610 GGGAGCGCGCGGTGGAGAGAGCCACCGAGAGAGCGGAGAGAGAGAGAGAGAG 668
Db 3110 CACGAGCTGTGAGAGAGAGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 3168

RESULT 14

US-10-033-026-3
; Sequence 3, Application US/10033026
; Patent No. US20020147309A1
; GENERAL INFORMATION:
; APPLICANT: Lipscombe, Diane
; APPLICANT: Schorge, Stephanie
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/10/033,026
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 09/268,163
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/077,901
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 7376
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 146..7174
US-10-033-026-3

Query Match 3.8%; Score 49; DB 12; Length 7376;
Best Local Similarity 48.0%; Pred. No. 0.016;

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGTGAAGTGAACCGCGCGGAGCGCTCGGTGACGTTTCACTCGGGCC 309
Db 2753 GAGCGCCGAGAGAGAGCGGGAGCGCGGTGCGCGGAGAGAGCGCGCGCGCAC 2812
QY 310 CGCAGCTGACGCTGACAGGTGCGGAGCGGACCGCGCGCGCGCGCGCGCGTCTCTCAA 369
Db 2813 CGCA-GCCACAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2869
QY 370 CCGCGTGGCCACCGCGCGCGCGGAGATCCCGCGATCTGCGCAGACCGTAGCCCGTTC 429
Db 2870 GGGCCAGCG 2929
QY 430 TCGTCCGTGACCTTCTGTGGCTCTCTCTCTCACTGAGGTGCGGAGAGCGAGAGACA 489
Db 2930 GAGCGGAGCG 2989
QY 490 CCCAGAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
Db 2990 GCGCGCAAGGCGAG 3049
QY 550 GTGGCCCGGAGGAAGCGGTCCCGACCGCGGAGCGAGAGGGCGATGTTCAAGCTGA 609
Db 3050 GCGGAGAGCGGGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3109
QY 610 GGGAGCGCGCGGTGGAGAGAGCGGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 668
Db 3110 CACGAGCTGTGAGAGAGAGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 3168

RESULT 15

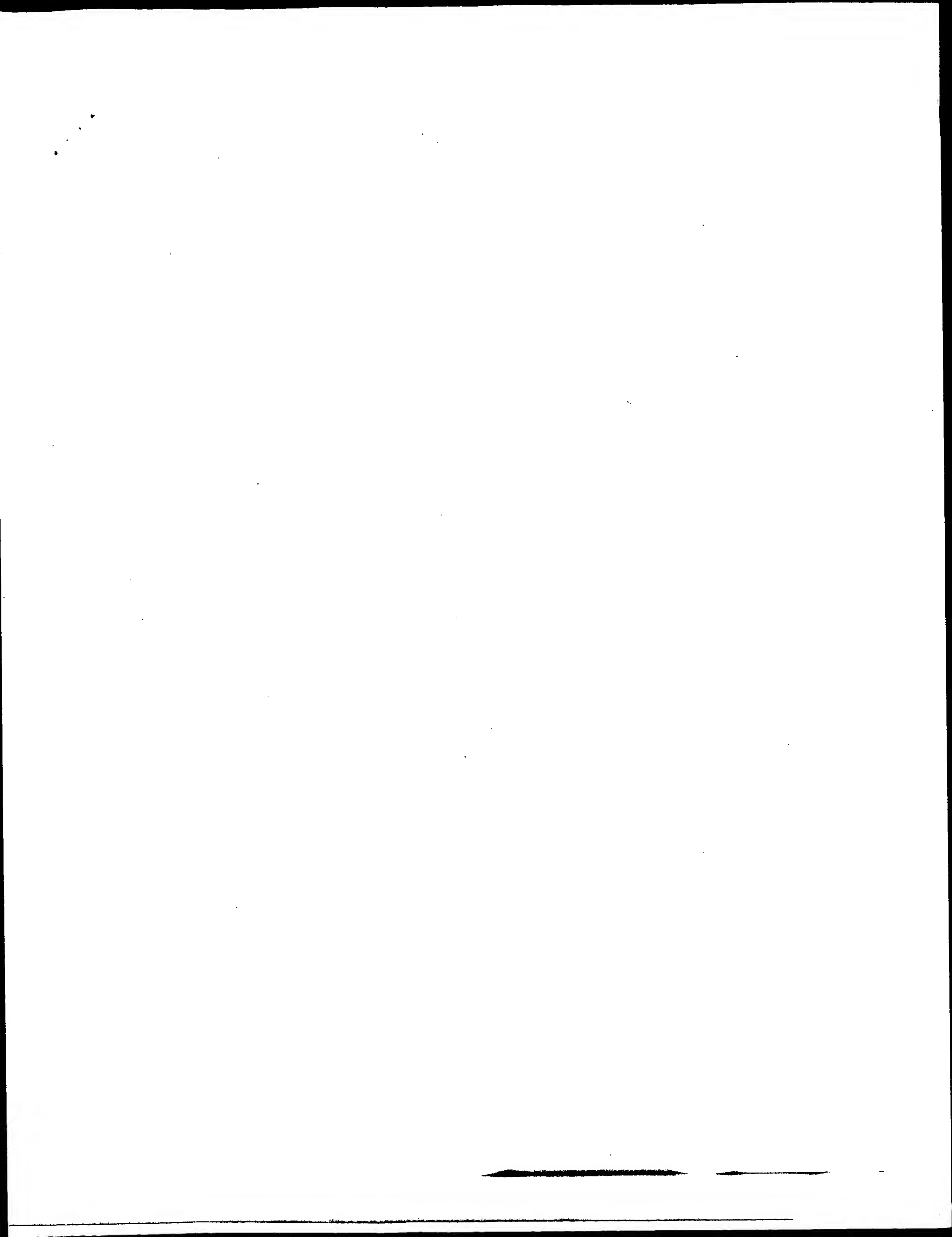
US-09-946-807-104/c
; Sequence 104, Application US/09946807
; Patent No. US20020165144A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinhorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/946,807
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-946-807-104

Query Match 3.6%; Score 46.6; DB 9; Length 1006;
Best Local Similarity 44.9%; Pred. No. 0.027;

Matches 217; Conservative 0; Mismatches 264; Indels 2; Gaps 1;

QY 178 GTGACAGCCACACGAGCGGACAGCTCATGCCCCCTGCTGCGGATGGTCCCGCTCG 237
Db 756 GTGACCTTACACAGATAGGCGCGCTCTCCCGGCGCTGCGGCGCGCGCGCGCG 699
QY 238 GTACAGACCGGTGACGCTGCGGTGACAGTGAACCCCGCGCGCGCGCGCGCGCG 297
Db 698 GCGGTGGCCAGAGAGGACCGGTCCCGTTGGCGGAGAGAGCGCGCTCTGCGCGCG 639
QY 298 TCATCGGGCG 357
Db 638 CCCAGCG 579
QY 358 GGTCTGTCAACCCCGTGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 417
Db 578 CCG 519
QY 418 GTAGCCCGCTTCTGCTGCTGACCTTCTGTGCGCTCTCTCTCACTGAGGTTGCGGA 477
Db 518 TTTCCTCGATCACACCGCGCGCGCGCGCTGAGCTAGCTCTCCACCGATCCACGCTGG 459
QY 478 GGCAGGACAGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537
Db 458 GGGACRAGTAGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
QY 538 GAGCCGAGAGAGTGGCGCGCGAGGAAGCGGTCCCGACCGCGGAGCGAGAGAGAG 597
Db 398 GCGGCGAGAGCGAG 339
QY 598 GTTACGCTGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 657
Db 338 GAGCGGCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 279
QY 658 GTG 660
Db 278 GCG 276

Search completed: January 19, 2003, 20:04:19
Job time : 88 secs



OM nucleic - nucleic search, using sw model

Run on: January 18, 2003, 09:43:41 ; Search time 296 Seconds
(without alignments)
9715.547 Million cell updates/sec

```

Title: US-09-844-864-1
Perfect score: 1277
Sequence: 1 aagcgcgcgagcgcgga.....acaaaaaaaaaaaaaaaaa 1277

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478
```

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : N_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	1265	99.1	1276	21	AAD00294	Mouse oocyte-specific
2	58.8	4.6	14955	20	AAK53491	Human adenosine A1
3	52	4.1	14955	20	AAK53491	Human adenosine A1
4	49.6	3.9	400	19	AAV64545	M. tuberculosis im
5	49.6	3.9	400	19	AAV44436	Mycobacterium tube
6	49.6	3.9	400	20	AAZ19346	M. tuberculosis an
7	49.6	3.9	400	20	AAZ19134	M. tuberculosis re
8	49.6	3.9	14800	24	ABL66291	Lung cancer relate
9	49.6	3.9	14835	24	AAS94858	Human DNA sequence

10	49	3.8	7175	16	AAQ84658	Human neuronal cal
11	49	3.8	7175	19	AAV42686	DNA encoding human
12	49	3.8	7175	21	AAA71704	Human calcium chan
13	49	3.8	7266	19	AAV29059	Human calcium chan
14	49	3.8	7362	14	AAQ37817	Sequence encoding
15	49	3.8	7362	16	AAQ84657	Human neuronal cal
16	49	3.8	7362	19	AAV42685	DNA encoding human
17	49	3.8	7362	21	AAA71703	Human calcium chan
18	49	3.8	7364	24	ABL65869	Lung cancer relate
19	49	3.8	7376	20	AAH88001	Human calcium chan
20	49	3.8	109519	22	AAH88001	N-type calcium cha
21	48.2	3.8	524	22	AAH88001	Micromonospora DNA
22	48.2	3.8	524	22	AAH88001	Human CDNA 5'-end
23	48.2	3.8	524	22	AAH88001	Human CDNA 5'-end
24	47.6	3.7	5944	24	ABL92203	Human CDNA clone r
25	47.6	3.7	5944	24	ABL32253	Human full-length
26	46.6	3.6	745	15	AAQ62837	Chemically treated
27	46.6	3.6	745	15	AAQ58330	Human immune syste
28	46.6	3.6	745	17	AAT48077	GGF gene coding se
29	46.6	3.6	745	17	AAT30988	GGF gene coding se
30	46.6	3.6	745	17	AAT30988	Human glial growth
31	46.6	3.6	745	20	AAZ32060	Human glial growth
32	46.6	3.6	745	20	AAZ32060	Human METH2 relate
33	46.6	3.6	745	22	AAH81175	Human neurogulin G
34	46.6	3.6	1006	22	AAH81175	I36448 CDNA clone.
35	46.6	3.6	1006	22	AAH81175	Human NRGI exon sl
36	46.6	3.6	1986	20	AAZ32061	Human NRGI exon sl
37	46.6	3.6	1986	22	AAZ32061	Human METH2 relate
38	46.6	3.6	2003	15	AAQ62838	L12260 CDNA clone.
39	46.6	3.6	2003	15	AAQ62838	GGF2HBS5. Homo sa
40	46.6	3.6	2003	17	AAQ62838	GGF2HBS5. Homo sa
41	46.6	3.6	2003	17	AAT48090	Human neurogulin G
42	46.6	3.6	2003	17	AAT30995	Glial growth facto
43	46.6	3.6	2003	20	AAT06739	Glial growth facto
44	46.6	3.6	2003	22	AAZ32062	Human METH2 relate
45	45.8	3.6	6232	13	AAQ29269	I36352 CDNA clone.
						Human calcium chan

ALIGNMENTS

Application
Priority

RESULT 1
AAD00294
AAD00294 standard; cDNA; 1276 BP

AAD00294

05-SEP-2000 (first entry)

Mouse oocyte-specific 01-180 cDNA clone.

Oocyte-specific; ovary; O1-180; mouse; gynaecological; treatment; screen; cell proliferative disorder; cell degenerative disorder; contraceptive; modulator; signalling pathway; human infertility; cancer; ovulation; ss.

Mus sp.

Key	Location/Qualifiers
CDS	28..1113

/product="Mouse oocyte-specific protein, 01-180"

MO200024755-AT

04-MAY-2000.

28-OCT-1999; 99WO-US25209

28-OCT-1998; 98US-0106020.

(BAYU) BAYLOR COLLEGE MEDICINE

Matzok MM, Wang P;

XX WP1: 2000-350684/30.
DR P-PSDB; AAY70948.
DR
XX
PT 01-180, 01-184 and 01-236 polypeptides and nucleic acids encoding them,
PT useful for evaluating potential contraceptives to block ovulation in a
PT reversible manner -
XX
PS Claim 2; Fig 1; 54pp; English.
XX
XX The present sequence is the cDNA encoding the mouse oocyte-specific
CC protein 01-180, expressed in the oocytes of primary (one-layer)
CC preantral follicles through ovulation. It provides in vitro and in vivo
CC reagents for studying ovarian development and function. This sequence
CC has gynaecological and contraceptive activity. Agents which modulate
CC 01-180, 01-184 and 01-236 may be used to treat cell proliferative or
CC degenerative disorders, associated with abnormal expression of these
CC ovary specific genes. This ovary-specific sequence can be used as
CC reagents to evaluate potential contraceptives, to block ovulation in a
CC reversible manner. It is also used to screen for genetic mutations in
CC signalling pathways, that are associated with some forms of human
CC infertility or gynaecological cancers.

Sequence 1276 BP; 279 A; 389 C; 394 G; 214 T; 0 other;

Query Match	99.1%;	Score 1265;	DB 21;	Length 1276;
Best Local Similarity	99.9%;	Pred. No. 2.3e-291;		
Matches 1276; Conservative	0;	Mismatches 0;	Indels 1;	Gaps 1;

QY	1	AAGGCGGGCGAGGGCGGGAGCCACCATTGTTCCGGCGAGACAGCTTCCACCCTGCCCC	60
Db	1	AAGGCGGGCGAGGGCGGGAGCCACCATTGTTCCGGCGAGACAGCTTCCACCCTGCCCC	60
QY	61	CATCCTTATCCGAGGGCCACCAAGCCGGGATGGCTGGAGTTCCGAGCCAGGGGCTGC	120
Db	61	CATCCTTATCCGAGGGCCACCAAGCCGGGATGGCTGGAGTTCCGAGCCAGGGGCTGC	120
QY	121	CGACCCGCGCCCCCTCCTTCCTCCCGGCTACAGACAGCTCATGGCCGGAGTACGTC	180
Db	121	CGACCCGCGCCCCCTCCTTCCTCCCGGCTACAGACAGCTCATGGCCGGAGTACGTC	180
QY	181	GACAGCCACCAGCGGGACAGCTCATGGCCCTGCTGTGCGCGGATGGTCCCGGTCGTC	240
Db	181	GACAGCCACCAGCGGGACAGCTCATGGCCCTGCTGTGCGCGGATGGTCCCGGTCGTC	240
QY	241	AGCAGCCGTGACGCTGCGGTGACGTGAACCCGCGCGGACGCGCTCGGTGACGTGTTCA	300
Db	241	AGCAGCCGTGACGCTGCGGTGACGTGAACCCGCGCGGACGCGCTCGGTGACGTGTTCA	300
QY	301	CTCGGCGCGCCACGCTGCAGCCTGCAGGGTGCCGAGCCAGCCCGCAGCCCGATCGGGT	360
Db	301	CTCGGCGCGCCACGCTGCAGCCTGCAGGGTGCCGAGCCAGCCCGCAGCCCGATCGGGT	360
QY	361	TCCTGTCAACCCCGTGCCACAGCGCGCGCGGGAGATCCCGCGCATCTGCAGACCGTA	420
Db	361	TCCTGTCAACCCCGTGCCACAGCGCGCGCGGGAGATCCCGCGCATCTGCAGACCGTA	420
QY	421	GCCCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCCTCACTGAGGTTCCGGAGGC	480
Db	421	GCCCCGTTCTCGTCCGTGACCTTCTGTGGCTCTCCTCACTGAGGTTCCGGAGGC	480
QY	481	AGGCAGACACCCACGAAGGAGAGGGAGCCCGGCATCCTCGGGACCCGGGAACCGGAG	540
Db	481	AGGCAGACACCCACGAAGGAGAGGGAGCCCGGCATCCTCGGGACCCGGGAACCGGAG	540
QY	541	CCGAGAGAGGTGGCCCGCAGGAAGCGGTCCCCACGCCGGAAGCGAGAGGGCGATGTT	600
Db	541	CCGAGAGAGGTGGCCCGCAGGAAGCGGTCCCCACGCCGGAAGCGAGAGGGCGATGTT	600
QY	601	CAGGCTGCAGGGGACGGCCGGGTGGGAGCAGCAGCCACCAACCGAGGACCGGAACAGTGTG	660
Db	601	CAGGCTGCAGGGGACGGCCGGGTGGGAGCAGCAGCCACCAACCGAGGACCGGAACAGTGTG	660

QY	661	GC	CGCGATG	CAGTTC	AGCCTG	GAGCGAGG	AGCCATG	TCTG	CCCGCAG	ATGG	CTCAG	720
Db	661	GC	GGCGATG	CAGTTC	AGCCTG	GAGCGAGG	AGCCATG	TCTG	CCCGCAG	ATGG	CTCAG	720
QY	721	GAC	CCCGTG	ATTGGA	TGCCCC	TCGAGAC	CAGGCTC	CCCGCA	AAGCAG	CGAGCAG		
Db	721	GAC	CCCGTG	ATTGGA	TGCCCC	TCGAGAC	CAGGCTC	CCCGCA	AAGCAG	CGAGCAG		
QY	781	AAG	GAGCGC	TGCGTT	CCAGTT	CTTAGAG	CAGAGT	ACGGCT	ACTAT	CACTG	CAAGGAC	840
Db	781	AAG	GAGCGC	TGCGTT	CCAGTT	CTTAGAG	CAGAGT	ACGGCT	ACTAT	CACTG	CAAGGAC	840
QY	841	TG	CAAAAT	CCGGTG	GAGAGC	GCCTAT	GTGTG	TGTGTG	CAGGGC	ACAGT	AAAGT	900
Db	841	TG	CAAAAT	CCGGTG	GAGAGC	GCCTAT	GTGTG	TGTGTG	CAGGGC	ACAGT	AAAGT	900
QY	901	CTT	CAACAG	TCTG	CCGAGT	GTGTG	AGAATC	CTCAAC	CCCTT	ACAGAG	TGAGG	960
Db	900	CTT	CAACAG	TCTG	CCGAGT	GTGTG	AGAATC	CTCAAC	CCCTT	ACAGAG	TGAGG	960
QY	961	CAC	CTGTCA	AGTTGT	TAAGA	ACTAG	ATGTG	CTG	CCAGT	CAGATTT	CGCCAC	1020
Db	960	CAC	CTGTCA	AGTTGT	TAAGA	ACTAG	ATGTG	CTG	CCAGT	CAGATTT	CGCCAC	1020
QY	1021	CC	CTAAAC	CGCCCC	ATCGG	CAGACT	TGTGT	GAGATG	CAAGCA	AAAGCG	CTGCTG	1080
Db	1020	CC	CTAAAC	CGCCCC	ATCGG	CAGACT	TGTGT	GAGATG	CAAGCA	AAAGCG	CTGCTG	1080
QY	1081	CG	CAGAC	CACCTT	CAGCTT	CAATAT	ATCAT	TATT	AGT	GAGAGT	CGAAAA	1140
Db	1080	CG	CAGAC	CACCTT	CAGCTT	CAATAT	ATCAT	TATT	AGT	GAGAGT	CGAAAA	1140
QY	1141	TG	GGGCTA	ATGGA	TGGACA	AGTGA	AGCTT	TCTCCC	CTTC	ACCTCT	TCCCTT	1200
Db	1140	TG	GGGCTA	ATGGA	TGGACA	AGTGA	AGCTT	TCTCCC	CTTC	ACCTCT	TCCCTT	1200
QY	1201	TCTT	CATGA	CAGAC	AGTGT	TACTTG	ATAT	AAAGC	CTGTGA	TAAAG	GTAT	1260
Db	1200	TCTT	CATGA	CAGAC	AGTGT	TACTTG	ATAT	AAAGC	CTGTGA	TAAAG	GTAT	1260
QY	1261	AAAA	AAAAAAAA	AAAAAA	1277							
Db	1260	AAAA	AAAAAAAA	AAAAAA	1276							

RESULT 2
AAx53491
ID AAX53491 standard; DNA; 114955 BP.

AC AAX53491;

DT 05-JUL-1999 (first entry)

Human adenosine A1 receptor antisense oligonucleotide fragment.

AA Antisense oligonucleotide; multiple target; antisense treatment;
 KW impaired respiration; inflammation; lung disease;
 KW pulmonary vasoconstriction; inflammation; allergic rhinitis;
 KW acute asthma; allergy; asthma; impeded respiration;
 KW respiratory distress syndrome; pain; cystic fibrosis;
 KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;
 KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
 KW colon cancer; breast cancer; lung cancer; pancreatic cancer;
 KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
 KW prostate cancer; ss.

Synthetic.

PN W09913886-A1.

PD 25-MAR-1999.

17-SEP-1998; 98WO-US19419.

XX 09-JUN-1998; 98US-0093972.
PR 17-SEP-1997; 97US-0059160.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PT Nyce JW;
XX WPI; 1999-229400/19.
XX
PT New antisense oligonucleotides used in treatment of, e.g. pulmonary
PT vasoconstriction
XX
PS Disclosure; Page 37; 120pp; English.
XX
CC The specification describes antisense oligonucleotides (AA52869-X55271)
CC directed against at least 2 mRNAs selected from target genes, coding and
CC non-coding regions of RNAs corresponding to target genes, coding and
CC initiation codons, genomic flanking regions, intron-exon borders, the
CC 5'-end, the 3'-end and the junction between coding and non-coding
CC regions and all segments of RNAs encoding proteins associated with one
CC or more diseases, conditions or mixtures. The antisense oligonucleotides
CC may be derived from sequences AAX55272-74. These multiple target
CC oligonucleotides (specifically AAX55180-271) can be used for the
CC antisense treatment of diseases and conditions. Typical diseases and
CC conditions are those associated with impaired respiration and
CC inflammation, including lung diseases, pulmonary vasoconstriction,
CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded
CC respiration, respiratory distress syndrome, pain, cystic fibrosis,
CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
CC obstructive pulmonary disease (COPD), and cancers such as leukemias,
CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,
CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,
CC hepatic metastases, as well as all types of cancers which may metastasize
CC or have metastasized to the lungs, including breast and prostate cancer.
XX
SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;
Query Match 4.6%; Score 58.8; DB 20; Length 114955;
Best Local Similarity 32.6%; Pred. No. 0.0013;
Matches 251; Conservative 71; Mismatches 441; Indels 8; Gaps 3;
QY 3 GGGGGGCGAGCGCGGAGCGACCGCCATGTTCCCGGAGACAGTTCACCCCTGCGCGCA 62
DB 104641 GGGCBGGGCGCGCGGNNNDNNCCGCBGGCGCGCGGCGGCGGCGGCGGCGGCGG 104700
QY 63 TCCTTATCCGAGGCGACCAAGCGCGGATGCTGAGGTTCGAGCGAGCGGCGTCCG 122
DB 104701 GC---GCGGNNNDNNCCGCBGGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 104756
QY 123 ACCCGCGCGCGCT 182
DB 104757 NDNCCGCBGGCGCGGCGGNNNDNNCCGCBGGCGGCGGCGGCGGCGGCGGCGG 104816
QY 183 CAGCCACGAGCGGCGACAGCTCATGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
DB 104817 SNNDNNCCGCBGGCGGNNNDNNCCGCBGGCGGCGGCGGCGGCGGCGGCGG 104876
QY 243 CAGCGTGACGCTGCGGTGAGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 104877 NCCGCBGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104936
QY 301 CTCGGGCGCGCGAGCGTGCAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 104937 CCBGGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 104996
QY 361 TCCTGTCAACCCCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 104997 GSNNDNNBGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 105056
QY 421 GCGCGCTTCT 480
DB 105057 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105116

QY 481 AGCAGACACCCACGAG 540
DB 105117 CCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105174
QY 541 CCGAGAGAGGTGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 105175 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105234
QY 601 CAGGCTGAGGCGGAG 660
DB 105235 CCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105294
QY 661 GCGGCGGATGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 105295 CCGGCGGCGGNNNDNNCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105354
QY 721 GACCGCGGTGATTGCGATGCGGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 771
DB 105355 NDNCCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 105405
RESULT 3
AAX53491/c
ID AAX53491 standard; DNA; 114955 BP.
XX
AC AAX53491;
XX
DT 05-JUL-1999 (first entry)
XX
DE Human adenosine A1 receptor antisense oligonucleotide fragment.
XX
KW Antisense oligonucleotide; multiple target; antisense treatment;
KW impaired respiration; inflammation; lung disease;
KW pulmonary vasoconstriction; inflammation; allergic rhinitis;
KW acute asthma; allergy; asthma; impeded respiration;
KW respiratory distress syndrome; pain; cystic fibrosis;
KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;
KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
KW colon cancer; breast cancer; lung cancer; pancreatic cancer;
KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
KW prostate cancer; ss.
XX
OS Synthetic.
XX
PN W09913886-A1.
XX
PD 25-MAR-1999.
XX
PF 17-SEP-1998; 98WO-US19419.
XX
PR 09-JUN-1998; 98US-0093972.
PR 17-SEP-1997; 97US-0059160.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX WPI; 1999-229400/19.
XX
PT New antisense oligonucleotides used in treatment of, e.g. pulmonary
PT vasoconstriction
XX
PS Disclosure; Page 37; 120pp; English.
XX
CC The specification describes antisense oligonucleotides (AAX52869-X55271)
CC directed against at least 2 mRNAs selected from target genes, coding and
CC non-coding regions of RNAs corresponding to target genes, coding and
CC initiation codons, genomic flanking regions, intron-exon borders, the
CC 5'-end, the 3'-end and the junction between coding and non-coding
CC regions and all segments of RNAs encoding proteins associated with one
CC or more diseases, conditions or mixtures. The antisense oligonucleotides
CC may be derived from sequences AAX55272-74. These multiple target

[illegible]

Bd 359 CGCGGGCAGGTCACCTCTCAGG 382

RESULT 6
AAZ19346
ID AAZ19346 standard; DNA: 400 BP.
XX
AC AAZ19346;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis antigen 5' XP22 DNA sequence.
XX
KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
immunotherapy; diagnosis; immunisation; vaccine; infection;
immune response; skin test; ss.
XX
OS Mycobacterium tuberculosis.
PN WO9942076-A2.
PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US03268.
XX
PR 05-MAY-1998; 98US-0072967.
PR 18-FEB-1998; 98US-0025197.
XX
PA (CORI-) CORIXA CORP.
PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skelky YAW, Twardzik DR, Vedvick TS;
XX WPI: 1999-527409/44.

New antigens from Mycobacterium tuberculosis useful in diagnostic
skin tests and protective or therapeutic vaccines or compositions

Claim 4; Page 170; 299pp; English.

The present invention describes polypeptides comprising an immunogenic
part of a Mycobacterium tuberculosis antigen (Ag). Also described
are vaccines and fusion protein containing M. tuberculosis Ag's.
M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
other polypeptides fragments, can be used in pharmaceutical compoitions
or vaccines to generate a protective or therapeutic immune response to
M. tuberculosis and as reagents in skin tests for diagnosis of
tuberculosis. Ag can induce proliferation of, or cytokine secretion
by, T, B or natural killer cells and/or macrophages in
tuberculosis-immune subjects. AAZ19249 to AAZ19460 and AAY39083 to
AAY39225 are used in the exemplification of the present invention.

Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;

Query Match 3.9%; Score 49.6; DB 20; Length 400;
Best Local Similarity 49.2%; Pred. No. 0.033;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 458 CCTCACTGAGGTTCGCGGAGCAGCACACCACGAAGGAGGGAGCCGCAT 517
| | | | |
Db 119 CTTCACCCAAGGCCGCGAGCAACGCCGCAACGGCGTGACGGCGGTCGGCGCAA 178
| | | | |
QY 518 CCTCGGGGACCCGGAACCGGAGCCGAGAAGAGTGGCCCCGAGAAAGCGTTCCCAGC 577
| | | | |
Db 179 CGGCGGAACAAGCGCGGAAGCGCGCAGACAACACACCACCCGCCGCCCGGCACACAGG 238
| | | | |
QY 578 CGCGAAGCAGAGAGGCGCATGTTCAAGCTGCAGAGGCGCAGGCCGGTGGAGCAGACCCAC 637
| | | | |
Db 239 CGGCGAAGCGCGCGCGCGCGCGCGCGCGGAACCGCGGGAACCGCGGAGCGCGCGCAC 298
| | | | |
QY 638 CACCGAGAGCAGCAGACTGTGCGCGCGCATGCAGTCTGAGCCTGGAGCGAGAGCCAT 697
| | | | |
Db 299 CGGCGACCGCGCGCGCGCGCGCGCGCGCAACCGCGCGCAACCGCGCGCAAGCGCGCAC 358

QY 698 GTCTGCGCGACAGATGGCTCAGG 721
| | | | |
Db 359 CGCGCGGACGGGTGCACCTCTCAGG 382

RESULT 7

AAZ19134
ID AAZ19134 standard; DNA; 400 BP.

XX AAZ19134;

DT 05-NOV-1999 (first entry)

DE M. tuberculosis recombinant antigen DNA encoding 5' XP22.

KW Antigen; diagnosis; detection; infection; antibody; immunisation;
KW vaccine; immunity; ss.

OS Mycobacterium tuberculosis.

PN WO9942118-A2.

PD 26-AUG-1999.

PF 17-FEB-1999; 99WO-US03265.

PR 05-MAY-1998; 98US-0072596.

PR 18-FEB-1998; 98US-0024753.

PA (CORI-) CORIXA CORP.

PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;

DR WPI; 1999-527416/44.

PT New polypeptide comprising antigenic portions of M. tuberculosis

PS Claim 4; Page 215-216; 323pp; English.

CC This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against
CC M. tuberculosis infection. The new detection methods are needed as
CC current vaccination strategies do not provide 100% immunity.

XX Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;

QY Query Match 3.9%; Score 49.6; DB 20; Length 400;

Best Local Similarity 49.2%; Pred. No. 0.033;
Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 458 CCTCACTGAGGTTCGGGAGGAGGACGACACCCAGAAAGGAGAGAGCCCGGCAT 517
| | | | |

Db 119 CTTACCCCAAGGCGCGACGCGCAACGCGCGGTGACGGCGGTGCGGCGCAA 178
| | | | |

QY 518 CCTCGGGAGCCCGGAAACCGAGCCGAGAGAGTGGCCGCGAGAAAGCGTCCCGCAGC 577
| | | | |

Db 179 CGCGGAAACGCGGAAACGCGCAGACACACACCGCCCGCGGACACACAGG 238
| | | | |

QY 578 CGCGAAGCAGAGAGGCGGATGTTCAAGCTGACAGGCGAGCCGCGTGGAGCAGACGCAC 637
| | | | |

Db 239 CGCGCAGCGCGCGCGCGCGCGCGGACCGCGGACCGCGCGCGCGCGCGCAC 298
| | | | |

QY 638 CACCGGAGGAGCCGGAACAGTGTGGCGCGGATGAGCTTGAGAGGAGAGGAGCAT 697
| | | | |

Db 299 CGGCACCGCAC 358
| | | | |

QY 698 GTCTGCGCGACAGATGGCTCAGG 721
| | | | |

Db 359 CGCGCGGACGGGTGCACCTCTCAGG 382

RESULT 8

ABL66291
ID ABL66291 standard; DNA; 14800 BP.

XX ABL66291;

DT 15-MAY-2002 (first entry)

DE Lung cancer related gene sequence SEQ ID NO:4628.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

XX gene; ds.

OS Homo sapiens.

PN WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US10838.

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.

PR 25-SEP-2000; 2000US-235077P.

PR 25-SEP-2000; 2000US-235082P.

PR 25-SEP-2000; 2000US-235134P.

PR 25-SEP-2000; 2000US-235280P.

PR 26-SEP-2000; 2000US-235637P.

PR 26-SEP-2000; 2000US-235638P.

PR 27-SEP-2000; 2000US-235711P.

PR 27-SEP-2000; 2000US-235720P.

PR 27-SEP-2000; 2000US-235840P.

PR 27-SEP-2000; 2000US-235863P.

PR 28-SEP-2000; 2000US-236028P.

PR 28-SEP-2000; 2000US-236032P.

PR 28-SEP-2000; 2000US-236033P.

PR 28-SEP-2000; 2000US-236034P.

PR 28-SEP-2000; 2000US-236109P.

PR 28-SEP-2000; 2000US-236111P.

PR 29-SEP-2000; 2000US-236842P.

PR 29-SEP-2000; 2000US-236891P.

PR 02-OCT-2000; 2000US-237172P.

PR 02-OCT-2000; 2000US-237173P.

PR 02-OCT-2000; 2000US-237278P.

PR 02-OCT-2000; 2000US-237294P.

PR 02-OCT-2000; 2000US-237295P.

PR 02-OCT-2000; 2000US-237316P.

PR 03-OCT-2000; 2000US-237425P.

PR 03-OCT-2000; 2000US-237598P.

PR 03-OCT-2000; 2000US-237604P.

PR 03-OCT-2000; 2000US-237606P.

PR 03-OCT-2000; 2000US-237608P.

PR 01-NOV-2000; 2000US-244867P.

PR 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrihan S;

PI Soppet DR, Weaver Z;

XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
XX determining a change in expression of a gene of a signature gene set -
PS Claim 1; SEQ ID 4628; 44pp; English.

XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 14800 BP; 2890 A; 4567 C; 5319 G; 2024 T; 0 other;

Query Match 3.9%; Score 49.6; DB 24; Length 14800;
Best Local Similarity 48.3%; Pred. No. 0.1;
Matches 139; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 507 GAGCCCGGCATCTCGGGGACCCGGGAACCGAGAGAGGTGGCCGAGAAAGC 566
DB 4122 GAGCCAGTACATCAAGTTTCATCAGCGAGACTCTGCGCGCATGAGAGAGAGGCT 4181
QY 567 GGTCCCCAGCCCGAAGCGAGAGGGCGATGCTGAGGCGAGCCGGTGGGA 626
DB 4182 GCGTAGCAGCAGCGGGGAGAGAGCGCGCGCTGGCCGAGTGGAGCCGCGTGA 4241
QY 627 GCAGCAGCCACCAACCGAGAGCGGAACAGTGTGGCGGATGAGTGGAGCTGGAG 686
DB 4242 GAAGCAGCGGCGAGTGGCGGAGCGCGCCAGCAAGGACACAGCGGAGCGGAGGC 4301
QY 687 CGAGAGCCATGCTCTGCGCAGAGATGCTCAAGACCCCGGTGATTCGATGCCCTCG 746
DB 4302 GAAGGAGCTGCAGCAGCGCATGACAGAGAGGTGTGCGGCGGAGAGCGCGGTGA 4361
QY 747 AGACCAGGCTTCCCGCAAAAGCAGCAGAGAGCAAGAGCGCTGCG 794
DB 4362 CGCGCAGCAGCAGAAAGCGCAGCATTCAGAGAGAGCTGCACAGACTGCG 4409

RESULT 9
AAS94858
ID AAS94858 standard; DNA; 14835 BP.

XX AAS94858;

DT 14-FEB-2002 (first entry)

DE Human DNA sequence #113 expressed during foam cell differentiation.

KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;
cardiovascular disorder; coronary artery disease; gene therapy; ds.

OS Homo sapiens.

PN WO200177389-A2.

PD 18-OCT-2001.

XX

PF 04-APR-2001; 2001WO-US11128.

PR 05-APR-2000; 2000US-195106P.

PA (INCY-) INCYTE GENOMICS INC.

PI Shiffman D, Somogyi R, Lawn R, Selhammer JJ, Porter GJ, Mikita T;
PI Tai J;

DR WPI; 2002-010925/01.

PT Composition useful for diagnosis of conditions, disorders or diseases
PT associated with atherosclerosis, comprises several polynucleotides that
PT are differentially expressed in foam cell development -

PS Claim 1; Page 165-169; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used
CC as PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. Aas94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation.

XX Sequence 14835 BP; 2881 A; 4568 C; 5306 G; 2018 T; 62 other;

Query Match 3.9%; Score 49.6; DB 24; Length 14835;
Best Local Similarity 48.3%; Pred. No. 0.1;
Matches 139; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 507 GAGCCCGGCATCTCGGGGACCCGGGAACCGAGAGAGGTGGCCGAGAAAGC 566
DB 4149 GAGCCAGTACATCAAGTTTCATCAGCGAGACTCTGCGCGCATGAGAGAGAGAGGCT 4208
QY 567 GGTCCCCAGCCCGAAGCGAGAGGGCGATGCTGAGGCGAGCCGGTGGGA 626
DB 4209 GCGTAGCAGCAGCGGGGAGAGAGCGCGCGCTGGCCGAGTGGAGCGCGCGTGA 4268
QY 627 GCAGCAGCCACCAACCGAGAGCGGAACAGTGTGGCGGATGAGTGGAGCTGGAG 686
DB 4269 GAAGCAGCGGCGAGCTGGCGGAGCGCACGCCAGCAAAAGCACAGCGCGGAGAGGC 4328
QY 687 CGAGAGCCATGCTCTGCGCAGAGATGCTCAGAGACCCCGGTGATTCGATGCCCTCG 746
DB 4329 GAAGGAGCTGCAGCAGCGCATGACAGAGAGAGTGTGCGGCGGAGAGAGCGCGGTGA 4388
QY 747 AGACCAGGCTTCCCGCAAAAGCAGCAGAGAGCAAGAGCGCTGCG 794
DB 4389 CGCGCAGCAGCAGAAAGCGCAGCATTCAGAGAGAGCTGCACAGACTGCG 4436

RESULT 10
AAQ84658
ID AAQ84658 standard; DNA; 7175 BP.

XX AAQ84658;

DT 01-DEC-1995 (first entry)

DE Human neuronal calcium channel subunit alpha 1B-2.

KW Calcium channel subunit; antagonist; agonist; diagnosis;
Lambert Eaton Syndrome; ss.

OS Homo sapiens.

XX

```

FH Key      Location/Qualifiers
FT CDS      144..6857
FT          /*tag= a
FT misc_feature 6633..7175
FT          /*tag= b
FT          /note= "identical to alpha 1B-1"
XX
XX WO9504822-A.
XX
XX 16-FEB-1995.
XX
XX 11-AUG-1994; 94WO-US09230.
XX
XX 11-AUG-1993; 93US-0105536.
XX
XX 05-NOV-1993; 93US-0149097.
XX
XX (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.
XX
XX Ellis SB, Gillespie A, Harpold MM, McCue AF, Williams ME;
XX
XX WPI; 1995-090900/12.
XX
XX P-PSDB; AAR71006.
XX
XX DNA encoding human calcium channel sub-unit(s) - used for
XX developing prods. for studying calcium channels, e.g. for
XX obtaining agonists and antagonists
XX
XX PS Disclosure; Page 149-160; 285pp; English.
XX
XX CC DNA encoding the alpha 1B subunit was isolated by screening a
XX human basal ganglia cDNA library with fragments of the rabbit
XX skeletal muscle calcium channel alpha 1 subunit-encoding cDNA.
XX A portion of one of the positive clones was used to screen an
XX IMR32 cell cDNA library. Clones that hybridised to the basal
XX ganglia probe were used to further screen an IMR32 cell cDNA
XX library to identify overlapping clones that in turn were used
XX to screen a human hippocampus cDNA library. A series of clones
XX to span nearly the entire length of the nt. sequence encoding
XX the human alpha 1B subunit was obtd. Nucleic acid amplification
XX of specific regions of the IMR32 cell alpha 1B mRNA yielded
XX additional segments of the alpha 1B coding sequence. A full-
XX length alpha 1B DNA clone was constructed by ligating portions
XX of the partial cDNA clones. Nucleic acid amplification analysis
XX of IMR32 cell RNA and genomic DNA using oligo primers corresp. to
XX sequences located 5' and 3' of the stop codon of the DNA encoding
XX the alpha 1B subunit revealed an alternatively spliced alpha
XX 1B-encoding mRNA in IMR32 cells. This second mRNA product is the
XX result of differential splicing of the alpha 1B subunit transcript
XX to include another exon that is not present in the mRNA corresp.
XX to the other 3' alpha 1B cDNA sequence that was initially isolated.
XX The alpha 1B subunit encoded by a DNA sequence contg. an additional
XX exon is referred to as alpha 1B-1 and given in AA084657/R71005,
XX whereas the other form is referred to as alpha 1B-2 and is given in
XX AA084658/R71006. Following the sequence of the additional exon in
XX alpha 1B-1 the alpha 1B-1 and alpha 1B-2 sequences are identical.
XX
XX SQ Sequence 7175 BP; 1415 A; 2197 C; 2168 G; 1395 T; 0 other;
XX
XX Query Match      3.8%; Score 49; DB 16; Length 7175;
XX Best Local Similarity 48.0%; Pred. No. 0.11;
XX Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;
XX
XX QY 250 GACGCTGCGGTGAGTGAACCGCGCGGAGCGCTCGGTGAGTGTTCACCTCGGGCGC 309
XX      ||||| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 2751 GAGGCCCCGAGAGCGGAGAGCGGCGGTGCGCGGAGAGAGCGCGCGCGCAC 2810
XX
XX QY 310 CGCAGCGTGCAGCTGCAAGGTTCCGAGCCAGCCCGGAGCGCGGATCGGTCTCTCAA 369
XX      ||||| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 2811 GCGA-GCCACAGCAGAGAGCGCGCGGGGCC--CCCGAGAGCGCGGAGCGCGCGCA 2867
XX
XX QY 370 CCCCCTGGCCACGCGCGCGCGGAGATCCCGCGGATCTGGCAGACCGTACCCCGTTC 429
XX      ||||| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 2868 GCGCCAGGCGCGCGCGCGCGGCGGACACCGCGCGGCTCCCGGAGAGAGCGCGCC 2927

```

```

QY 430 TCGTCCGTACCTTCTGTGCGCTTCCTCCTCCTACTGAGGTTGCGGAGGACGACAGACA 489
XX      ||||| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 2928 GAGCGGAGAGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2987
XX
XX QY 490 CCCACGAGAGAGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 549
XX      ||||| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 2988 GCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3047
XX
XX QY 550 GTGGCCGCGAGAAAGCGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
XX      ||||| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 3048 GCGGAGAGCGCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3107
XX
XX QY 610 GGGCAGCGCGCGGTGCGAGCAGCAGCCACCGGAGGACCGGAGCGCGCGCGCGAT 668
XX      ||||| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 3108 CACGAGGCTGTGAGAAAGAGACACCGAGAGAGAGAGCGCGCGCGCGCGCGCGAT 3166
XX
XX RESULT 11
XX AAV42686
XX ID AAV42686 standard; DNA; 7175 BP.
XX
XX AC AAV42686;
XX
XX 12-OCT-1998 (first entry)
XX
XX DE DNA encoding human calcium channel alpha-1B-1 subunit.
XX
XX KW Alpha-1B subunit; human; calcium channel; assay; detection;
XX characterisation; Lambert Eaton Syndrome; LES; diagnosis; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key      Location/Qualifiers
XX FT 5'UTR    1..143
XX FT CDS      144..6857
XX FT          /*tag= a
XX FT          /*tag= b
XX FT 3'UTR    6855..7175
XX FT          /*tag= c
XX
XX PN US5792846-A.
XX
XX PD 11-AUG-1998.
XX
XX PE 31-MAY-1995; 95US-0455543.
XX
XX PR 04-APR-1994; 94US-0223305.
XX PR 04-APR-1988; 88US-0176899.
XX PR 04-APR-1989; 89US-0603751.
XX PR 04-APR-1989; 89WO-US01408.
XX PR 20-FEB-1990; 90US-0482384.
XX PR 30-NOV-1990; 90US-0620250.
XX PR 15-AUG-1991; 91US-0745206.
XX PR 31-MAY-1995; 95US-0455543.
XX
XX PA (SIBI-) SIBIA NEUROSCIENCES INC.
XX
XX PI Brenner R, Ellis SB, Feldman DH, Harpold MM, McCue AF;
XX PI Williams ME;
XX
XX DR WPI; 1998-456192/39.
XX DR P-PSDB; AAW63142.
XX
XX PT DNA encoding human calcium channel alpha 1B sub:unit protein -
XX PT useful for recombinant production of the channel for screening of
XX PT its modulators, and diagnosis of Lambert Eaton Syndrome
XX
XX PS Claim 1; Columns 91-106; 166pp; English.
XX
XX CC The present sequence encodes the alpha-1B subunit of a human calcium
XX CC channel. The present sequence is derived from alternative splicing of
XX CC AAV42685. Calcium channels are membrane-spanning, multi-subunit proteins

```


CC that allow controlled entry of calcium ions into cells. This leads
CC to depolarisation events required for muscle contraction. The
CC recombinant subunit, when expressed with nucleic acids encoding the
CC complete calcium channel, can be used in assays for the detection and
CC characterisation of compounds that modulate the channel. The DNA encoding
CC the subunits can be alternatively spliced when transcribed, giving more
CC than one form of the protein from the same transcript, each having
CC slightly different properties. In addition, the reactivity of the alpha 1
CC subunit with IgG molecules from the serum of an individual with Lambert
CC Eaton Syndrome (LES) can be used as a diagnostic for the disease.
XX
SQ Sequence 7175 BP; 1415 A; 2197 C; 2168 G; 1395 T; 0 other;

Query Match 3.8%; Score 49; DB 19; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.11;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGGTGACGTGACCCCGCGCGACGCTCGGTGACGTTCACCTCGGGCC 309
DB 2751 GAGGCCCCGAAGGCGGAGAGCGGAGCCCGGTGCGCGGAGAGCGCGCGCGCCAC 2810
QY 310 CGCAGCTGACGCTGACGCGGTGCGGACCGCCGACGCGCGATCGGTCTGTCAA 369
DB 2811 CGCA-GCCACAGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2867
QY 370 CCCGCTGGCCACGCGCGCGCGGAGATCCCGCGATCTGCGACAGCTAGCCCGTTTC 429
DB 2868 GCGCCAGGCG 2927
QY 430 TCGTCCGTGACCTTCTGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGAGAGCG 2987
QY 490 CCCACGAAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
DB 2988 GCGCGCAAGGCG 3047
QY 550 GTGGCCGCGAGGAACCGGTCTCCCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 609
DB 3048 GCGGAGAGCG 3107
QY 610 GCGCAGGCGCGGTGGAGCAGCAGCCACCGAGAGCGCGCGCGCGCGCGCGCGCGAT 668
DB 3108 CACGAGGCTGTGAGAGAGAGCAGCAGCGAGAGAGCGCGCGCGCGCGCGCGCGAT 3166

RESULT 12
ID AAA71704 standard; DNA: 7175 BP.

AC AAA71704;
DT 22-DEC-2000 (first entry)

XX Human calcium channel alpha1B-2 subunit encoding DNA.
DE Human calcium channel alpha1B-2 subunit encoding DNA.
XX
KW Human; calcium channel; calcium channel subunit; diagnosis;
KW Lambert Eaton Syndrome; calcium channel subunit alpha1B-2; ds.
XX
OS Homo sapiens.

FT Key Location/Qualifiers
FT CDS 144..6857
FT /tag= a
FT /product= "calcium channel alpha1B-2 subunit"

PN US6096514-A.
XX
PD 01-AUG-2000.
XX
PF 25-MAY-1995; 95US-0450562.
XX
PR 04-APR-1988; 88US-0176899.

PR 02-FEB-1990; 90US-0482384.
PR 08-NOV-1990; 90US-0603751.
PR 30-NOV-1990; 90US-0620250.
PR 15-AUG-1991; 91US-0745206.
PR 10-APR-1992; 92US-0868354.
PR 13-JUL-1992; 92US-0914231.
PR 11-AUG-1993; 93US-0105536.
PR 05-NOV-1993; 93US-0149097.
PR 07-FEB-1994; 94US-0193078.
PR 04-APR-1994; 94US-0223305.
PR 11-AUG-1994; 94US-0290012.
PR 23-SEP-1994; 94US-0311363.
PR 28-SEP-1994; 94US-0314083.
PR 07-NOV-1994; 94US-0336257.
PR 13-MAR-1995; 95US-0404950.

PA (SIBI-) SIBIA NEUROSCIENCES INC.
XX
XX Ellis SB, Williams ME, McCue AF, Harpold MM;
XX
XX WPI; 2000-548230/50.
XX P-PSDB; AAB10573.

PT Human calcium channel beta subunit polynucleotides, useful for
PT producing recombinant eukaryotic cells and for diagnosing Lambert Eaton
PT Syndrome -
XX
XX Example II; Column 113-128; 153pp; English.

CC This invention describes a novel isolated DNA molecule (I) comprising a
CC sequence encoding a beta3-1 subunit of a human calcium channel.
CC Nucleic acid probes comprising 14-30 contiguous nucleotides of
CC beta-3 subunit encoding DNA are useful for isolation and cloning of
CC calcium channel subunit-encoding DNA. Recombinant eukaryotic cells that
CC express heterologous calcium channel are useful for identifying compounds
CC that modulate calcium channel activity and in assays for identifying
CC agonists and antagonists of calcium channel activity in humans. Human
CC calcium channel subunit or eukaryotic cells expressing the channel are
CC useful for diagnosing Lambert Eaton Syndrome (LES) in a human. This
CC sequence encodes the human calcium channel alpha1B-2 subunit which is
CC described in the method of the invention.

SQ Sequence 7175 BP; 1415 A; 2197 C; 2168 G; 1395 T; 0 other;

Query Match 3.8%; Score 49; DB 21; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.11;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGCGGTGACGTGACCCCGCGCGACGCTCGGTGACGTTCACCTCGGGCC 309
DB 2751 GAGGCCCCGAAGGCGGAGAGCGGAGCCCGGTGCGCGGAGAGCGCGCGCGCCAC 2810
QY 310 CGCAGCTGACGCTGACGCGGTGCGGACCGCCGACGCGCGATCGGTCTGTCAA 369
DB 2811 CGCA-GCCACAGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2867
QY 370 CCCGCTGGCCACGCGCGCGCGGAGATCCCGCGATCTGCGACAGCTAGCCCGTTTC 429
DB 2868 GCGCCAGGCG 2927
QY 430 TCGTCCGTGACCTTCTGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGAGAGCG 2987
QY 490 CCCACGAAGGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 549
DB 2988 GCGGCAAGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3047
QY 550 GTGGCCGCGAGGAACCGGTCTCCCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 609
DB 3048 GCGGAGAGCG 3107
QY 610 GCGCAGGCGCGGTGGAGCAGCAGCCACCGAGAGCGCGCGCGCGCGCGCGCGAT 668

CC to identify overlapping clones that in turn were used to screen a
CC human hippocampus cDNA library. In this way, a sufficient series of
CC clones to span nearly the entire length of the nucleotide sequence
CC encoding the human alpha 1B subunit was obtained. PCR amplification
CC of specific regions of the IMR32 cell alpha 1B mRNA yielded
CC additional segments of the alpha 1B coding sequence. A full-length
CC alpha 1B DNA clone was constructed by ligating portions of the
CC partial cDNA clones (see AAQ37817, AAQ37818). Alpha 1B-1 and alpha
CC 1B-2 are derived by alternative splicing of the alpha 1B subunit
transcript.

XX Sequence 7362 BP; 1446 A; 2276 C; 2214 G; 1426 T; 0 other;

Query Match 3.8%; Score 49; DB 14; Length 7362;
Best Local Similarity 48.0%; Pred. No. 0.12;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

OY 250 GACGCTGGCGTGCAGGTGAACCCGCGCCGCGACCCCTCGGTGCACTGTTCACTCGGGCGC 309
DB 2751 GAGGCCCCGAAAGCGAGAGCGGGGAGCCCGGTGCGCCCGGAGAGCGCGCGCGCCAC 2810
OY 310 CGCAGCTGCAGCCTGCAGGTGCGGAGCCAGCCCGGACCGCCGATCGGTTCTGTCAA 369
DB 2811 CGCA-GCCACAGCAAGAGCGCGCGCGCGCC--CCCGAGGCGCGGAGCGAGCGCGCGCA 2867
OY 370 CCGCGTGGCCACGCGCGCGCGGAGATCCCGCGATCTGGCAGACCGTAGCCCGTTC 429
DB 2868 GCGCCAGGCG 2927
OY 430 TCGTCCGTGACCTTCTGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGAGAGCG 2987
OY 490 CCCACGAAGGAGAGGAGCG 549
DB 2988 GCGCGCAAGGAGGAGCG 3047
OY 550 GTGGCCCGGAGAAAGCGGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
DB 3048 GCGGAGAGGCG 3107
OY 610 GCGCAGGCGCGGTGCGAGCAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 668
DB 3108 CACGAGGCTGTGAGAGAGGAGCAGCGAGAGAGGAGCGCGCGCGCGCGCGCGCGCG 3166

RESULT 15

AAQ84657
ID AAQ84657 standard; DNA: 7362 BP.

XX AAQ84657;
AC
XX
DT 01-DEC-1995 (first entry)

DE Human neuronal calcium channel subunit alpha 1B-1.

XX
KW Calcium channel subunit; antagonist; agonist; diagnosis;
KW Lambert Eaton Syndrome; ss.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT CDS 144..7163

FT misc_difference 6633..6819
FT misc_feature 6820..7362
FT misc_feature 6820..7362
FT misc_feature 6820..7362

XX
XX /note= "Identical to alpha 1B-2"

PN WO9504822-A.

PD 16-FEB-1995.
XX
XX 11-AUG-1994; 94WO-US09230.
XX
XX 11-AUG-1993; 93US-0105536.
XX
XX 05-NOV-1993; 93US-0149097.
XX
XX (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
XX
XX Ellis SB, Gillespie A, Harpold MM, McCue AF, Williams ME;
XX
XX WPI; 1995-090900/12.
XX
XX P-PSDB; AAR71005.

PT DNA encoding human calcium channel sub-unit(s) - used for
PT developing prods. for studying calcium channels, e.g. for
PT obtaining agonists and antagonists
PS
XX
XX Disclosure: Page 138-149; 285pp; English.

CC DNA encoding the alpha 1B subunit was isolated by screening a
CC human basal ganglia cDNA library with fragments of the rabbit
CC skeletal muscle calcium channel alpha 1 subunit-encoding cDNA.
CC A portion of one of the positive clones was used to screen an
CC IMR32 cell cDNA library. Clones that hybridised to the basal
CC ganglia probe were used to further screen an IMR32 cell cDNA
CC library to identify overlapping clones that in turn were used
CC to screen a human hippocampus cDNA library. A series of clones
CC to span nearly the entire length of the nt. sequence encoding
CC the human alpha 1B subunit was obtained. Nucleic acid amplification
CC of specific regions of the IMR32 cell alpha 1B mRNA yielded
CC additional segments of the alpha 1B coding sequence. A full-
CC length alpha 1B DNA clone was constructed by ligating portions
CC of the partial cDNA clones. Nucleic acid amplification analysis
CC of IMR32 cell RNA and genomic DNA using oligo primers corresp. to
CC sequences located 5' and 3' of the stop codon of the DNA encoding
CC the alpha 1B subunit revealed an alternatively spliced alpha
CC 1B-encoding mRNA in IMR32 cells. This second mRNA product is the
CC result of differential splicing of the alpha 1B subunit transcript
CC to include another exon that is not present in the mRNA corresp.
CC to the other 3' alpha 1B cDNA sequence that was initially isolated.
CC The alpha 1B subunit encoded by a DNA sequence contg. an additional
CC exon is referred to as alpha 1B-1 and given in AAQ84657/R71005,
CC whereas the other form is referred to as alpha 1B-2 and is given in
CC AAQ84658/R71006. Following the sequence of the additional exon in
CC alpha 1B-1 the alpha 1B-1 and alpha 1B-2 sequences are identical.

XX
SQ Sequence 7362 BP; 1445 A; 2278 C; 2214 G; 1425 T; 0 other;

Query Match 3.8%; Score 49; DB 16; Length 7362;
Best Local Similarity 48.0%; Pred. No. 0.12;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

OY 250 GACGCTGGCGTGCAGGTGAACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 309
DB 2751 GAGGCCCCGAAAGCGAGAGCGGGGAGCCCGGTGCGCCCGGAGAGCGCGCGCGCGCA 2810
OY 310 CGCAGCTGCAGCCTGCAGGTGCGGAGCCAGCCCGGACCGCCGATCGGTTCTGTCAA 369
DB 2811 CGCA-GCCACAGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2867
OY 370 CCGCGTGGCCACGCGCGCGCGCGGAGATCCCGCGATCTGTGCAAGCGTAGCCCGTTC 429
DB 2868 GCGCCAGGCG 2927
OY 430 TCGTCCGTGACCTTCTGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 489
DB 2928 GAGCGGAGAGCG 2987
OY 490 CCCACGAAGGAGAGGAGCG 549
DB 2988 GCGCGCAAGGAGGAGCG 3047

Mon Jan 20 09:05:59 2003

us-09-844-864-1.rng

Page 12

QY 550 GTGGCCGCGAGGAAGCGGTCTCCCCAGCCGCCGAGAAGGAGGCGGCATGTTCAGCCTGCA 609
+
Db 3048 GCGAGAGCGGGGAGGAGCCGCGCGCGCACCGGGCCCGGCACAAAGGCCGAGCCTGCT 3107

QY 610 GGCAGAGCCGGGTGGAGCAGCAGCCACCACCGGAGGACCGGAACAGTGTGGCGGCAT 668
+
Db 3108 CACGAGCTGTGAGAAGAGAGACCAACGAGAGGAGGCCACGGAGAGAGGCGCTGAGAT 3166

Search completed: January 19, 2003, 10:25:45
Job time : 508 secs